

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平10-179941

(43)公開日 平成10年(1998) 7月 7日

(51)Int.Cl.<sup>6</sup>

識別記号

F I

A 6 3 H 3/33

A 6 3 H 3/33

C

5/00

5/00

C

G 1 0 L 3/00

G 1 0 L 3/00

Q

5 5 1

5 5 1 H

5 5 1 M

審査請求 有 請求項の数22 O L (全 31 頁)

(21)出願番号 特願平9-262557

(22)出願日 平成 9 年(1997) 9 月26日

(31)優先権主張番号 特願平8-278353

(32)優先日 平 8 (1996)10月21日

(33)優先権主張国 日本 (J P)

(71)出願人 596151548

桑原 淳司

東京都江東区永代 1-12-5 第7小崎ビル  
504

(71)出願人 596151559

高柳 尚

東京都新宿区矢来町115東海神楽坂マンシ  
ョン801

(72)発明者 桑原 淳司

東京都江東区永代 1-12-5 第7小崎ビル  
504

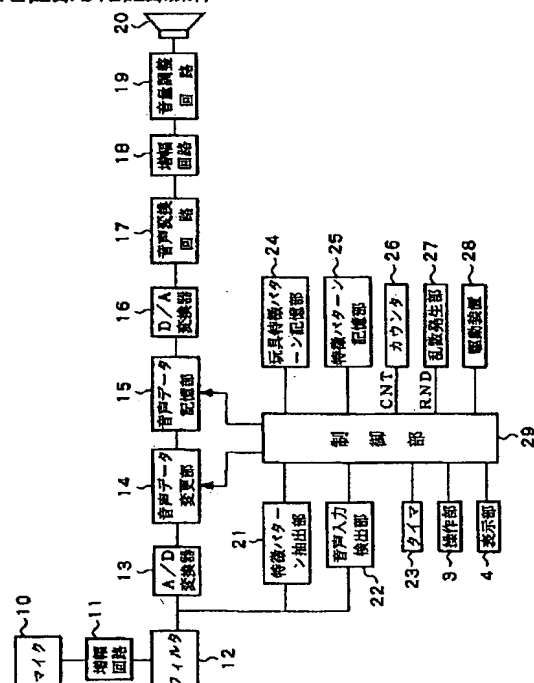
(74)代理人 弁理士 志賀 正武 (外11名)

(54)【発明の名称】 音声認識および音声発生装置、および、該音声認識および音声発生装置を備えた玩具、ならびに、音声認識および音声発生制御プログラムを記録した記録媒体

(57)【要約】

【課題】 入力された音声と、その音声に対して発声する音声との対応を任意に設定できる音声認識および音声発生装置、およびその装置を備えた玩具ならびに音声認識および音声発生制御プログラムを記録した記録媒体を提供すること。

【解決手段】 マイク10から入力された音は、A/D変換器13にて音声データに変換され、最終的に音声データ記憶部15に記憶される。また、特徴パターン抽出部21では別途マイク10から入力された音声の特徴パターンが抽出され、抽出された特徴パターンは、操作部3からの指示に基づいて音声データ記憶部15に記憶されたいずれかの音声データと対応づけられた後、特徴パターン記憶部25に記憶される。これにより、特徴パターン記憶部25に記憶された特徴パターンが再度抽出された場合その特徴パターンに対応する音声データが音声データ記憶部15から読み出されて最終的にスピーカ20から音として発せられる。



## 【特許請求の範囲】

【請求項 1】 入力された音声に対応する音声信号を出力する音声信号出力手段と、

前記音声信号出力手段から出力された音声信号を記憶する音声信号記憶手段と、

前記音声信号出力手段から出力された音声信号に基づいて、前記音声信号出力手段に入力された音声認識する音声認識手段と、

前記音声信号記憶手段に記憶された音声信号と、前記音声認識手段により認識された音声との対応を設定する対応設定手段と、

前記音声認識手段により、前記対応づけられた音声認識されると、前記対応設定手段によって設定された対応に従って、その認識した音声に対応する音声信号を、前記音声信号記憶手段から読み出す音声信号読出手段と、前記音声信号読出手段が読み出した音声信号を音声に変換して出力する音声出力手段とを具備する音声認識および音声発生装置を、人形または動物の形状を有するぬいぐるみ、あるいは、室内飾りの、内部または外部に設けたことを特徴とする音声認識および音声発生装置を備えた玩具。

【請求項 2】 時刻を指定する時刻指定手段と、

現在時刻を計時する計時手段と、

前記計時手段が前記時刻指定手段によって指定された時刻を計時すると、前記音声信号記憶手段に記憶された音声信号を繰り返し読み出す反復読出手段と、

前記対応設定手段によって設定された対応に従って、前記反復読出手段によって繰り返し読み出された音声信号に対応する、特定の音声認識手段によって認識されると、前記反復読出手段による音声信号の読出を停止させる読出停止手段とを有してなり、前記音声出力手段が、前記反復読出手段によって読み出された音声信号を音声に変換して出力することを特徴とする請求項 1 に記載の音声認識および音声発生装置を備えた玩具。

【請求項 3】 前記計時手段によって計時された現在時刻を音声によって通知する現在時刻通知手段と、

前記音声認識手段によって特定の音声認識された時、前記現在時刻通知手段が現在時刻を通知するように制御する時刻通知制御手段とを具備することを特徴とする請求項 1 または 2 に記載の音声認識および音声発生装置を備えた玩具。

【請求項 4】 前記時刻通知制御手段は、

前記計時手段が、前記時刻指定手段によって指定された時刻を計時した時に、前記現在時刻通知手段が現在時刻を通知するように制御することを特徴とする請求項 3 に記載の音声認識および音声発生装置を備えた玩具。

【請求項 5】 前記音声信号記憶手段に記憶された音声信号のうち、いずれか 1 つを選択する選択手段と、

前記計時手段が、前記時刻指定手段によって指定された

時刻を計時した時に、前記選択手段によって選択された音声信号を、前記音声信号記憶手段から読み出す指定時刻読出手段とを有してなり、

前記音声出力手段が、前記指定時刻読出手段によって読み出された音声信号を音声に変換して出力することを特徴とする請求項 1 に記載の音声認識および音声発生装置を備えた玩具。

【請求項 6】 前記音声信号出力手段から出力された音声信号を前記音声信号記憶手段に任意に記憶させる記憶制御手段と、

前記音声信号記憶手段に任意に記憶された音声信号のうち、任意の音声信号を読み出す任意読出手段とを有し、前記音声出力手段は、前記任意読出手段によって読み出された、任意に記憶された音声信号を音声に変換して出力することを特徴とする請求項 1 ないし 5 のうちいずれか 1 項に記載の音声認識および音声発生装置を備えた玩具。

【請求項 7】 乱数を発生する乱数発生手段と、

前記乱数発生手段から発生された乱数が、予め設定された数値と一致しなかった場合、前記音声信号出力手段から出力される音声信号に含まれる情報を変更して出力する音声信号変更手段とを有してなり、

前記音声出力手段が、前記音声信号変更手段から出力された音声信号を音声に変換して出力することを特徴とする請求項 1 ないし 6 のうちいずれか 1 項に記載の音声認識および音声発生装置を備えた玩具。

【請求項 8】 前記音声出力手段から出力される音声の音質を変更する音質変更手段を具備することを特徴とする請求項 7 に記載の音声認識および音声発生装置を備えた玩具。

【請求項 9】 前記音声信号出力手段から出力された音声信号が、前記音質変更手段により音質を変更された音声信号であった場合、前記乱数発生手段が発生する乱数の数値範囲を変化させる乱数発生制御手段を具備することを特徴とする請求項 8 に記載の音声認識および音声発生装置を備えた玩具。

【請求項 10】 前記音声認識手段により、予め設定された回数以上、同一の音声認識された場合、前記乱数発生手段が発生する乱数の数値範囲を変化させる乱数発生制御手段を具備することを特徴とする請求項 8 に記載の音声認識および音声発生装置を備えた玩具。

【請求項 11】 所定の数値範囲で常時計数するカウンタと、

前記音声信号出力手段から音声信号が出力された時、前記カウンタの計数値を読み取り、その計数値が所定の数値と一致した場合は、前記乱数発生手段が発生する乱数の数値範囲を変化させる乱数発生制御手段を具備することを特徴とする請求項 8 に記載の音声認識および音声発生装置を備えた玩具。

【請求項 12】 所定の音声信号を記憶した所定音声信

号記憶手段と、

前記カウンタの計数値と、前記所定の数値とが一致した場合、前記所定音声信号記憶手段から前記所定の音声信号を読み出す所定音声信号読出手段とを有し、前記音声出力手段が、前記所定音声信号読出手段によって読み出された所定の音声信号を音声に変換して出力することを特徴とする請求項11に記載の音声認識および音声発生装置を備えた玩具。

【請求項13】 前記音声信号記憶手段に記憶された音声信号を順次全て読み出す順次読出手段を有してなり、前記音声出力手段が、前記順次読出手段によって読み出された音声信号を音声に変換して出力することを特徴とする請求項1ないし11のうち、いずれか1項に記載の音声認識および音声発生装置を備えた玩具。

【請求項14】 前記人形または動物の形状を有するぬいぐるみ、もしくは、室内飾りは可動部を有し、前記音声出力手段から音声出力された時に、その可動部を駆動させる駆動手段を具備することを特徴とする請求項1ないし12のうち、いずれか1項に記載の音声認識および音声発生装置を備えた玩具。

【請求項15】 入力された音声に対応する音声データを出力する音声データ出力手段と、前記音声データ出力手段から出力された音声データを記憶する音声データ記憶手段と、前記音声データ出力手段から出力された音声データに基づいて、前記音声データ出力手段に入力された音声認識する音声認識手段と、前記音声データ記憶手段に記憶された音声データと、前記音声認識手段により認識された音声との対応を設定する対応設定手段と、前記対応設定手段により対応づけられた音声音声認識手段により認識されると、前記対応設定手段によって設定された対応に従って、前記認識した音声に対応する音声データを、前記音声データ記憶手段から読み出す音声データ読出手段と、前記音声データ読出手段が読み出した音声データを音声に変換して出力する音声出力手段とを具備することを特徴とする、コンピュータにより音声認識および音声発生を行う音声認識および音声発生装置。

【請求項16】 人または動物の画像を表示する画像表示手段と、前記音声データ読出手段が前記音声データ記憶手段から音声データを読み出す時に、前記画像表示手段に表示された人または動物の画像を変化するように制御する画像制御手段を有することを特徴とする請求項15に記載のコンピュータにより音声認識および音声発生を行う音声認識および音声発生装置。

【請求項17】 入力された音声に対応する音声データを出力する音声データ出力手段と、音声に対応する音声データを音声に変換して出力する音声出力手段とを有す

るコンピュータに音声認識および音声発生を行わせるためのプログラムを記録した記録媒体であって、該プログラムは前記音声データ出力手段から出力された音声データを記憶させ、前記音声データ出力手段から出力された音声データに基づいて前記入力された音声認識させ、前記記憶させた音声データと、前記認識した音声とを対応づけさせ、

前記対応づけさせた音声再度認識された時、前記記憶させた音声データの中からその音声に対応する音声データを前記音声出力手段に出力させることを特徴とする音声認識および音声発生制御プログラムを記録した記録媒体。

【請求項18】 前記プログラムは、前記コンピュータに、人または動物の画像を画像表示装置に表示させ、前記記憶させた音声データの中からその音声に対応する音声データを前記音声出力手段に出力させる際、前記画像表示装置に表示させた人または動物の画像を変化させることを特徴とする請求項17に記載の音声認識および音声発生制御プログラムを記録した記録媒体。

【請求項19】 前記プログラムは、前記コンピュータに、前記音声データ出力手段から出力された音声データを、前記認識された音声に対応付けられる音声データとして記憶させる際に、前記音声データ出力手段から出力された音声データを一時記憶し、該一時記憶させた音声データと同一の音声データが前記音声データ出力手段から出力された時に乱数を発生させ、

該乱数と所定値との大小関係が所定の条件を満たす毎に、前記画像表示装置に表示させる人または動物の画像に、それぞれ異なる変化を生じさせ、前記所定の条件を所定回数満たした時に、当該音声データを前記認識された音声に対応付けられる音声データとして記憶させることを特徴とする請求項18に記載の音声認識および音声発生制御プログラムを記録した記録媒体。

【請求項20】 前記プログラムは、前記認識された音声に対応付けられる音声データとして音声データを記憶させる際に、当該音声データに対し、当該音声データを消去するか否かを決定するための記憶保持値を発生し、該記憶保持値を所定時間が経過する毎に減少させ、一定値以下になった時に当該音声データを消去することを特徴とする請求項18に記載の音声認識および音声発生制御プログラムを記録した記録媒体。

【請求項21】 前記プログラムは、前記コンピュータに、前記対応付けられる音声データとして記憶させた音声データと、前記認識した音声とを対応づけさせる際

前記対応付けられる音声データとして記憶させた音声データを所定時間間隔で順次前記音声出力手段に出力させ、  
前記対応付けられる音声データとして記憶させた音声データを前記所定時間間隔で音声出力手段に出力させている間に入力された音声を確認し、  
該認識した音声に対して、その直前に前記音声出力手段に出力させた音声データとの関連度を示す情報を発生させ、  
該認識した音声を、該認識した音声に対して発生された情報によって関連度が示される音声データが前記音声出力手段に出力された直後に、再度認識した場合、その情報が示す関連度を増加させ、  
該関連度が所定の条件を満たした時に、前記認識した音声と前記対応付けられる音声データとして記憶させた音声データとを対応付けることを特徴とする請求項19に記載の音声認識および音声発生制御プログラムを記録した記録媒体。

【請求項22】 前記プログラムは、所定時間が経過する毎に前記情報が示す関連度を減少させ、該関連度が一定値以下になった時に、対応付けられた音声と、音声データとの対応を解消することを特徴とする請求項21に記載の音声認識および音声発生制御プログラムを記録した記録媒体。

#### 【発明の詳細な説明】

#### 【0001】

【発明の属する技術分野】この発明は、入力された音声に対応する音声を発する音声認識および音声発生装置、および、該装置を備えた玩具、ならびに、音声認識および音声発生制御プログラムを記録した記録媒体に関する。

#### 【0002】

【従来の技術】従来より、入力された音声を記憶し、記憶された音声をそのまま発音すると共に、再生スイッチ等を備え、その再生スイッチを操作することによって、必要に応じて録音した音を再生する玩具が考案されている。例えば、特開昭63-290594号公報においては、オウムのぬいぐるみにマイクロフォン、AD変換器、半導体メモリ、DA変換器、および、スピーカ等を備え、マイクロフォンに向かって発生された音声をAD変換器等によってデジタル信号化して音声データとして半導体メモリに記録し、マイクロフォンからの音声信号がなくなると、半導体メモリから順次音声データを読み出してDA変換器に出力し、スピーカから音声として発生させている。また、上述した玩具においては、再生スイッチ等を設けて半導体メモリに記憶されていた音声データを、所望する時に再生させる機能を有している。

#### 【0003】

【発明が解決しようとする課題】ところで、上述した玩具においては、入力された音声は、常に正確に記憶され

発音されるので、何等意外性がなくすぐに飽きられてしまう可能性があった。また、入力された音声をそのまま発音するだけなので、単なる玩具や相手にメッセージを伝える録音再生装置としての用途に限られていた。

【0004】この発明は上述した事情に鑑みてなされたもので、外部から入力する音声と、その音声に対して発音する音声との対応を任意に設定できる音声認識および音声発生装置、および、その装置を備えた玩具、ならびに、音声認識および音声発生制御プログラムを記録した記録媒体を提供することを目的としている。また、他の目的としては、その用途が玩具や録音再生装置のみにとどまらず、所定の時刻に使用者が起こすべき行動の内容等を音声で通知するスケジュール通知機能や目覚まし機能を有し、様々な用途に対応可能な音声認識および音声発生装置、および、その装置を備えた玩具、ならびに、音声認識および音声発生制御プログラムを記録した記録媒体を提供することを目的としている。

#### 【0005】

【課題を解決するための手段】請求項1に記載の発明は、入力された音声に対応する音声信号を出力する音声信号出力手段と、前記音声信号出力手段から出力された音声信号を記憶する音声信号記憶手段と、前記音声信号出力手段から出力された音声信号に基づいて、前記音声信号出力手段に入力された音声を認識する音声認識手段と、前記音声信号記憶手段に記憶された音声信号と、前記音声認識手段により認識された音声との対応を設定する対応設定手段と、前記音声認識手段により、前記対応づけられた音声認識されると、前記対応設定手段によって設定された対応に従って、その認識した音声に対応する音声信号を、前記音声信号記憶手段から読み出す音声信号読出手段と、前記音声信号読出手段が読み出した音声信号を音声に変換して出力する音声出力手段とを具備する音声認識および音声発生装置を、人形または動物の形状を有するぬいぐるみ、あるいは、室内飾りの、内部または外部に設けたことを特徴とする音声認識および音声発生装置を備えた玩具である。

【0006】請求項2に記載の発明は、請求項1に記載の音声認識および音声発生装置を備えた玩具において、時刻を指定する時刻指定手段と、現在時刻を計時する計時手段と、前記計時手段が前記時刻指定手段によって指定された時刻を計時すると、前記音声信号記憶手段に記憶された音声信号を繰り返し読み出す反復読出手段と、前記対応設定手段によって設定された対応に従って、前記反復読出手段によって繰り返し読み出された音声信号に対応する、特定の音声の前記音声認識手段によって認識されると、前記反復読出手段による音声信号の読出を停止させる読出停止手段とを有してなり、前記音声出力手段が、前記反復読出手段によって読み出された音声信号を音声に変換して出力することを特徴としている。

【0007】請求項3に記載の発明は、請求項1または

2に記載の音声認識および音声発生装置を備えた玩具において、前記計時手段によって計時された現在時刻を音声によって通知する現在時刻通知手段と、前記音声認識手段によって特定の音声認識された時、前記現在時刻通知手段が現在時刻を通知するように制御する時刻通知制御手段とを具備することを特徴としている。

【0008】請求項4に記載の発明は、請求項3に記載の音声認識および音声発生装置を備えた玩具において、前記時刻通知制御手段は、前記計時手段が、前記時刻指定手段によって指定された時刻を計時した時に、前記現在時刻通知手段が現在時刻を通知するように制御することを特徴としている。

【0009】請求項5に記載の発明は、請求項2ないし4のうちいずれか1項に記載の音声認識および音声発生装置を備えた玩具において、前記音声信号記憶手段に記憶された音声信号のうち、いずれか1つを選択する選択手段と、前記計時手段が、前記時刻指定手段によって指定された時刻を計時した時に、前記選択手段によって選択された音声信号を、前記音声信号記憶手段から読み出す指定時刻読出手段とを有してなり、前記音声出力手段が、前記指定時刻読出手段によって読み出された音声信号を音声に変換して出力することを特徴としている。

【0010】請求項6に記載の発明は、請求項1ないし5のうちいずれか1項に記載の音声認識および音声発生装置を備えた玩具において、前記音声信号出力手段から出力された音声信号を前記音声信号記憶手段に任意に記憶させる記憶制御手段と、前記音声信号記憶手段に任意に記憶された音声信号のうち、任意の音声信号を読み出す任意読出手段とを有し、前記音声出力手段は、前記任意読出手段によって読み出された、任意に記憶された音声信号を音声に変換して出力することを特徴としている。

【0011】請求項7に記載の発明は、請求項1ないし6のうちいずれか1項に記載の音声認識および音声発生装置を備えた玩具において、乱数を発生する乱数発生手段と、前記乱数発生手段から発生された乱数が、予め設定された数値と一致しなかった場合、前記音声信号出力手段から出力される音声信号に含まれる情報を変更して出力する音声信号変更手段とを有してなり、前記音声出力手段が、前記音声信号変更手段から出力された音声信号を音声に変換して出力することを特徴としている。

【0012】請求項8に記載の発明は、請求項7に記載の音声認識および音声発生装置を備えた玩具において、前記音声出力手段から出力される音声の音質を変更する音質変更手段を具備することを特徴としている。

【0013】請求項9に記載の発明は、請求項8に記載の音声認識および音声発生装置を備えた玩具において、前記音声信号出力手段から出力された音声信号が、前記音質変更手段により音質を変更された音声信号であった場合、前記乱数発生手段が発生する乱数の数値範囲を変

化させる乱数発生制御手段を具備することを特徴としている。

【0014】請求項10に記載の発明は、請求項8に記載の音声認識および音声発生装置を備えた玩具において、前記音声信号出力手段から出力された音声信号が、前記音質変更手段により音質を変更された音声信号であった場合、前記乱数発生手段が発生する乱数の数値範囲を変化させる乱数発生制御手段を具備することを特徴としている。

10 【0015】請求項11に記載の発明は、請求項8に記載の音声認識および音声発生装置を備えた玩具において、所定の数値範囲で常時計数するフリーランカウンタと、前記音声信号出力手段から音声信号が出力された時、前記フリーランカウンタの計数値を読み取り、その計数値が所定の数値と一致した場合は、前記乱数発生手段が発生する乱数の数値範囲を変化させる乱数発生制御手段を具備することを特徴としている。

20 【0016】請求項12に記載の発明は、請求項11に記載の音声認識および音声発生装置を備えた玩具において、所定の音声信号を記憶した所定音声信号記憶手段と、前記フリーランカウンタの計数値と、前記所定の数値とが一致した場合、前記所定音声信号記憶手段から前記所定の音声信号を読み出す所定音声信号読出手段とを有し、前記音声出力手段が、前記所定音声信号読出手段によって読み出された所定の音声信号を音声に変換して出力することを特徴としている。

30 【0017】請求項13に記載の発明は、請求項1ないし12のうち、いずれか1項に記載の音声認識および音声発生装置を備えた玩具において、前記音声信号記憶手段に記憶された音声信号を順次全て読み出す順次読出手段を有してなり、前記音声出力手段が、前記順次読出手段によって読み出された音声信号を音声に変換して出力することを特徴としている。

40 【0018】請求項14に記載の発明は、請求項1ないし13のうち、いずれか1項に記載の音声認識および音声発生装置を備えた玩具において、前記人形または動物の形状を有するぬいぐるみ、もしくは、室内飾りは可動部を有し、前記音声出力手段から音声出力された時に、その可動部を駆動させる駆動手段を具備することを特徴としている。

50 【0019】請求項15に記載の発明は、入力された音声に対応する音声データを出力する音声データ出力手段と、前記音声データ出力手段から出力された音声データを記憶する音声データ記憶手段と、前記音声データ出力手段から出力された音声データに基づいて、前記音声データ出力手段に入力された音声を認識する音声認識手段と、前記音声データ記憶手段に記憶された音声データと、前記音声認識手段により認識された音声との対応を設定する対応設定手段と、前記対応設定手段により対応づけられた音声認識手段により認識されると、前

記対応設定手段によって設定された対応に従って、前記認識した音声に対応する音声データを、前記音声データ記憶手段から読み出す音声データ読出手段と、前記音声データ読出手段が読み出した音声データを音声に変換して出力する音声出力手段とを具備することを特徴とする、コンピュータにより音声認識および音声発生を行う音声認識および音声発生装置である。

【0020】請求項16に記載の発明は、請求項15に記載の音声認識および音声発生装置において、人または動物の画像を表示する画像表示手段と、前記音声データ読出手段が前記音声データ記憶手段から音声データを読み出す時に、前記画像表示手段に表示された人または動物の画像を変化するように制御する画像制御手段を有することを特徴としている。

【0021】請求項17に記載の発明は、入力された音声に対応する音声データを出力する音声データ出力手段と、音声に対応する音声データを音声に変換して出力する音声出力手段とを有するコンピュータに音声認識および音声発生を行わせるためのプログラムを記録した記録媒体であって、該プログラムは、前記コンピュータに前記音声データ出力手段から出力された音声データを記憶させ、前記音声データ出力手段から出力された音声データに基づいて前記入力された音声を確認させ、前記記憶させた音声データと、前記認識した音声とを対応づけさせ、前記対応づけさせた音声に再度認識された時、前記記憶させた音声データの中からその音声に対応する音声データを前記音声出力手段に出力させることを特徴としている。

【0022】請求項18に記載の発明は、請求項17に記載の音声認識および音声発生制御プログラムを記録した記録媒体において、前記コンピュータに、人または動物の画像を画像表示装置に表示させ、前記記憶させた音声データの中からその音声に対応する音声データを前記音声出力手段に出力させる際、前記画像表示装置に表示させた人または動物の画像を変化させるプログラムを記録していることを特徴としている。

【0023】請求項19に記載の発明は、請求項18に記載の音声認識および音声発生制御プログラムを記録した記録媒体において、前記プログラムが、前記コンピュータに、前記音声データ出力手段から出力された音声データを、前記認識された音声に対応付けられる音声データとして記憶させる際に、前記音声データ出力手段から出力された音声データを一時記憶し、該一時記憶させた音声データと同一の音声データが前記音声データ出力手段から出力された時に乱数を発生させ、該乱数と所定値との大小関係が所定の条件を満たす毎に、前記画像表示装置に表示させる人または動物の画像に、それぞれ異なる変化を生じさせ、前記所定の条件を所定回数満たした時に、当該音声データを前記認識された音声に対応付けられる音声データとして記憶させることを特徴としてい

る。

【0024】請求項20に記載の発明は、請求項18に記載の音声認識および音声発生制御プログラムを記録した記録媒体において、前記プログラムが、前記認識された音声に対応付けられる音声データとして音声データを記憶させる際に、当該音声データに対し、当該音声データを消去するか否かを決定するための記憶保持値を発生し、該記憶保持値を所定時間が経過する毎に減少させ、一定値以下になった時に当該音声データを消去することを特徴としている。

【0025】請求項21に記載の発明は、請求項19に記載の音声認識および音声発生制御プログラムを記録した記録媒体において、前記プログラムが、前記コンピュータに、前記対応付けられる音声データとして記憶させた音声データと、前記認識した音声とを対応づけさせる際に、前記対応付けられる音声データとして記憶させた音声データを所定時間間隔で順次前記音声出力手段に出力させ、前記対応付けられる音声データとして記憶させた音声データを前記所定時間間隔で音声出力手段に出力させている間に入力された音声を確認し、該認識した音声に対して、その直前に前記音声出力手段に出力させた音声データとの関連度を示す情報を発生させ、該認識した音声を、該認識した音声に対して発生された情報によって関連が示される音声データが前記音声出力手段に出力された直後に、再度認識した場合、その情報が示す関連度を増加させ、該関連度が所定の条件を満たした時に、前記認識した音声と前記対応付けられる音声データとして記憶させた音声データとを対応付けることを特徴としている。

【0026】請求項22に記載の発明は、請求項21に記載の音声認識および音声発生制御プログラムを記録した記録媒体において、前記プログラムが、所定時間が経過する毎に前記情報が示す関連度を減少させ、該関連度が一定値以下になった時に、対応付けられた音声と、音声データとの対応を解消することを特徴としている。

【0027】

【発明の実施の形態】以下、図面を参照して本発明の一実施形態について説明する。図1は、本実施形態における音声認識および音声発生装置を備えた玩具（以下、単に玩具という）の外観を示す図であり、この図において、1はインコを模したぬいぐるみであり、その頭部には入力された音声を電気信号に変換するマイクロフォン（図示略）が、また、喉部には音声を発するスピーカ（図示略）がそれぞれ内蔵されている。また、1aはインコを模したぬいぐるみの嘴であり、スピーカから音声が発せられる時に、あたかもぬいぐるみのインコがしゃべっているかのように開閉するようになっている。

【0028】2は、本実施形態における玩具の本体部であり、上述したマイクロフォンから音声が入力されると、その音声に応じた音声信号をスピーカへ出力する音

10

20

30

40

50

声信号発生装置（詳しくは後述する）と、この音声信号発生装置の電源である電池とが内蔵されている。3は、本体部2に取り付けられた複数の押しボタンからなる操作部であり、以下に示す各種のボタンにより構成されている。

#### 【0029】(1) モード選択ボタン

本実施形態における玩具の動作モードを選択するボタンであり、以下の各モードに対応してそれぞれ1つずつボタンが設けられている。

- ①学習モード…本実施形態における玩具に言葉または音を習得させるモード。
- ②設定モード…現在時刻、または、学習モードにおいて習得させた言葉もしくは音のうち、いずれかのものを、どの時刻に、または、どのような音もしくは音声、マイクroフォンに入力された時に発声させるか、その対応関係を設定するモード。
- ③実行モード…設定モードにおいて設定した内容を実行するモード。
- ④ランダムモード…入力された音声を任意に記憶し、また、本実施形態における玩具が記憶している言葉または音を任意に発声させるモード。
- ⑤メッセージモード…記憶した言葉または音を順次発声させるモード。

#### 【0030】(2) 時刻設定ボタン

設定モードにおいて時刻を設定する時に押下する（詳しくは後述する）。

#### (3) 現在時刻ボタン

設定モードにおいて、本実施形態の玩具に現在時刻を音声で通知させる時に押下する（詳しくは後述する）。

#### (4) 目覚ましボタン

本実施形態における玩具が有する目覚まし機能（詳しくは後述する）を有効または無効にするボタン。

#### (5) テンキー

0～9の数字に対応するボタンにより構成されている。

#### (6) 決定ボタン

上述した各モードにおいて、設定等を決定する際に押下される。さらに、4は表示部であり、上述した操作部3を構成する各種ボタンによって設定された内容等を表示する。

【0031】次に、前述した音声認識および音声発生装置の構成について、図2を参照して説明する。ここで、図1に示した各部に相当する部分については同一の符号を付し、その説明を省略する。この図において、10はマイクroフォン（以下、マイクという）であり、図1のぬいぐるみ1の頭部に内蔵され、入力された音声に応じた電気信号（以下、音声信号という）を出力する。11は増幅回路であり、マイク10から出力された音声信号を増幅する。12はフィルタであり、可聴帯域外の電気信号を除去する。13はA/D変換器であり、フィルタ12を通過した音声信号を所定周期でサンプリング化

し、デジタル信号に変換して音声データとして出力する。

【0032】14は音声データ変更部であり、A/D変換器13から出力された音声データを一時記憶すると共に、制御部29からの指示により、一時記憶した音声データを音声データ記憶部15またはD/A変換器16のいずれかに出力する。また、その際、制御部29からの指示に基づいて、出力する音声データの一部または全部を改竄する。音声データ記憶部15は、音声データ変更部14から出力された音声データを、制御部29により指定されたアドレスに格納する。D/A変換器16は、音声データ変更部14または音声データ記憶部15から出力された音声データをアナログ信号に変換する。

【0033】17は音声変換回路であり、実際のインコの音声に類似した音声となるように、D/A変換器16から出力されたアナログ信号の波形を適宜調節する。18は増幅回路であり、音声変換回路17から出力された音声信号を増幅する。19は音量調整回路であり、本実施形態における玩具が発する音声の音量を調節可能とするために、増幅回路18から出力された音声信号の振幅を調整する。20はスピーカであり、ぬいぐるみ1の喉部に内蔵され、音量調整回路19から出力された音声信号を音に変換して出力する。

【0034】21は特徴パターン抽出部であり、フィルタ12から出力された音声信号から線形予測係数や、フーリエスペクトル係数等を求め、マイク10に入力された音声の特徴パターンを抽出する。22は音声入力検出部であり、マイク10に入力された音声が所定レベル以上であった場合、検出信号を制御部29へ出力する。23はタイマであり、現在時刻を計時する。24は玩具特徴パターン記憶部であり、マイク10から入力された音声別の本実施形態における玩具から発せられたものか否かを判断するために、音声変換回路17によって変換される音声信号の特徴パターンが記憶されている。

【0035】25は特徴パターン記憶部であり、特徴パターン抽出部21から出力された特徴パターンを、制御部29により指示されたアドレスに記憶する。26はカウンタであり、本実施形態における玩具の電源がONされてからOFFされるまで、所定の数値範囲を常時計数する。ここで、カウンタ26は、0～99（10進数）の数値範囲内で常時計数している。27は乱数発生部であり、制御部29からの指示により、0～1の数値範囲内で、小数点第3位までの乱数RNDを出力する。28は駆動装置であり、制御部29からの指示により、ぬいぐるみ1の嘴1aを図示せぬモータ等によって開閉させる。29は制御部であり、上述した各部を制御する。

【0036】次に、図3ないし図8に示すフローチャートを参照して、制御部29の制御動作について説明する。まず、本実施形態の玩具の図示せぬ電源スイッチが



ONにされると、制御部29は、操作部3におけるモード選択ボタンの押下待機状態となる。そして、モード選択ボタンが押下された場合は、押下されたモードボタンに対応するモードの処理を開始する。以下、各種モード毎にそれぞれの処理について説明する。

【0037】〔学習モードにおける処理〕操作部3の学習モードボタンが押下されると、制御部29は、図3に示すフローチャートの処理を開始する。ここで、本実施形態における玩具から発生される音声は、基本的に外部から入力された音声であるが、本実施形態の玩具は、入力された音声を単にそのまま記憶するのではなく、所定の確率に従って入力された音声を正確に記憶するか否かを決定するものである。

【0038】図3に示すフローチャートの処理が開始されると、制御部29はステップS a 1へ進み、学習モードにおいて使用する以下のフラグをリセットする。

・玩具音声判定フラグMAC

マイク10に入力された音声は、本実施形態における玩具と同じ玩具から発生されたものであるか否かを示すフラグ。同じ玩具から発生されたものであれば「1」となり、同じ玩具から発生されたものでなければ「0」となる。

・リーチフラグREA

外部から入力された音声を記憶する確率が高くなるか否かを示すフラグ。記憶する確率を高くする場合は「1」となり、通常確率のままである場合は「0」となる。

【0039】ステップS a 1において、初期化が完了すると、ステップS a 2へ進み、制御部29は、マイク10から入力される音声の待機状態となる。すなわち、音声入力検出部22から検出信号が出力されたか否かを判断し、検出信号が出力されなかった場合は、マイク10から音声が入力されなかったとみなし、判断結果がNOとなって、再度ステップS a 2の処理を行う。そして、マイク10に音声が入力され、音声入力検出部22から検出信号が出力されると、判断結果はYESとなり、その後、再び検出信号が出力されなくなった時点で、ステップS a 3の処理へ進む。また、この間、音声データ変更部14には、A/D変換器13から出力された音声データが一時記憶される。

【0040】ステップS a 3において、制御部29は特徴パターン抽出部21から出力された特徴パターンと、玩具特徴パターン記憶部24に記憶されている特徴パターンとを比較する。そして、ステップS a 4へ進み、特徴パターン抽出部21から出力された特徴パターンと、玩具特徴パターン記憶部24に記憶されている特徴パラメータとを比較し、両者が類似していると判断した場合は、特徴パターンマイク10に入力された音声は、本実施形態の玩具と同じ玩具から発生されたものとみなし、判断結果がYESとなって、ステップS a 5へ進んで玩具音声判定フラグMACを「1」にセットした後、ステ

ップS a 6へ進む。また、両者が類似していないと判断された場合は、判断結果がNOとなり、ステップS a 5の処理を行わずに、ステップS a 6へ進む。

【0041】ステップS a 6において、制御部29は、カウンタ26のカウント値CNTを読み取る。そして、読み取った値が定数aであるか否かを判断する。ここで、定数aは0~49のうち、いずれか1つの整数であり、その値は予め定められている。そして、カウント値CNTが整数aと一致した場合は、判断結果がYESとなって、ステップS a 8へ進み、リーチフラグREAを「1」にセットした後、ステップS a 9にて、音声データ記憶部15に対し、特定の音声データが記憶されたアドレスを指定し、その音声データをD/A変換器16へ出力させる。

【0042】そして、音声データ記憶部15から出力された特定の音声データは、D/A変換器16にてアナログ音声信号に変換された後、音声変換回路17、増幅回路18、および、音量調整回路19を経て、スピーカ20から特定音（例えば、実際のインコ独特のさえずり等）として発音される。また、使用者は、この特定音を聞くことにより、今、入力した音声を、本実施形態の玩具が正確に記憶する確率が高くなったことを認識することができる。そして、制御部29は、ステップS a 9の処理が完了すると、ステップS a 10への処理へ進む。

【0043】一方、ステップS a 7において、カウント値CNTと整数aが一致しなかった場合は、ステップS a 8、S a 9の処理を行わずに、ステップS a 10へ進む。ステップS a 10では、特徴パターン抽出部21から出力された特徴パターンと、特徴パターン記憶部25のアドレスP(0)からP(i)までに記憶された特徴パターンとを順次比較していき、特徴パターン記憶部25に一致する特徴パターンがあるか否かを判断する。

【0044】そして、ステップS a 11へ進み、一致する特徴パターンがなかった場合は、判断結果がNOとなり、ステップS a 12へ進んで変数iの値を1インクリメントし、ステップS a 13にて特徴パターン抽出部21から出力された特徴パターンを、特徴パターン記憶部25のアドレスP(i)に記憶させ、ステップS a 15へ進む。一方、ステップS a 11において、特徴パターン記憶部25内に一致するパターンがあった場合は、判断結果がYESとなってステップS a 14へ進み、そのパターンが記憶されているアドレスP(i)のiの値を検出する。

【0045】次にステップS a 15へ進み、今回マイク10に入力された音声は、今までに何回入力されたのかをカウントするために、入力回数カウントデータC

(i)の値を1つインクリメントする。次いで、ステップS a 16へ進み、今回マイク10に入力された音声の入力回数が30回以上になったか否かを判断する。ここで、もし、入力回数が30回未満であった場合は判断結



果がNOとなり、ステップS a 17へ進む。

【0046】次いでステップS a 17では、玩具音声判定フラグMACまたはリーチフラグREAのいずれかが「1」にセットされているか否かを判断する。ここで、玩具音声判定フラグMACおよびリーチフラグREAのいずれも「1」にセットされていない場合は判断結果がNOとなり、ステップS a 18へ進む。そして、制御部29は、乱数発生部27に対して乱数RNDを出力させ、その乱数RNDに基づいて0～29の数値範囲内で整数の乱数を発生する。

【0047】次いで、ステップS a 19へ進み、ステップS a 18における処理により得られた乱数が、定数bと一致するか否かを判断する。ここで、定数bは0～29のうち、いずれか1つの整数であり、その値は予め定められている。そして、ステップS a 18にて得られた乱数と、定数bとが一致しなかった場合は、判断結果がNOとなってステップS a 20へ進み、音声データ変更部14に対し、一時記憶している今回入力された音声データの一部または全部を改竄するように指示する。これにより、音声データ変更部14は、上記音声データの一部または全部を、故意に異なるデータに変更する。

【0048】次に制御部29はステップS a 21の処理へ進み、音声データ変更部14に対して一時記憶している音声データをD/A変換器16へ出力させる。そして、D/A変換器16へ出力された音声データは、アナログ音声信号に変換された後、音声変換回路17においてインコのような音色に変換され、増幅回路18および音量調整回路19を経て、スピーカ20から音声が発せられる。また、この時、発せられる音声の内容は、マイク10に入力された音声の内容と異なっている。ここで、本実施形態においては、スピーカ20から音声が発せられている間に、音声入力検出部22から検出信号が出力された場合、すなわち、マイク10へ何らかの音が入力された場合、制御部29はその音を無効とみなすものとする。

【0049】一方、ステップS a 19において、ステップS a 18で得られた乱数と、定数bとが一致した場合は、ステップS a 20における処理を行わず、直接ステップS a 21の処理へ進む。これにより、マイク10から入力された音声の内容そのまま、しかもインコのような音色でスピーカ20から発せられる。

【0050】なお、ステップS a 16において、今回マイク10に入力された音声の入力回数が、30回以上であると判断された場合、または、ステップS a 17において、玩具音声判定フラグMACもしくはリーチフラグREAのいずれかが「1」にセットされていると判断された場合は、ステップS a 22へ進み、制御部29は、乱数発生部27に対して乱数RNDを出力させ、その乱数RNDに基づいて0～2の数値範囲内で整数の乱数を発生する。

【0051】そして、ステップS a 23へ進み、ステップS a 22における処理により得られた乱数が、定数cと一致するか否かを判断する。ここで、定数cは0～2のうち、いずれか1つの整数であり、その値は予め定められている。そして、ステップS a 22にて得られた乱数と、定数cとが一致しなかった場合は、判断結果がNOとなってステップS a 20へ進み、また、一致した場合は、ステップS a 21へ進み、以下、前述した処理と同じ処理を行う。

10 【0052】次いで、制御部29は、ステップS a 21の処理が終了すると、ステップS a 24へ進み、駆動装置27に対してぬいぐるみ1の嘴1aを駆動するよう指示する。これにより、駆動装置27は、スピーカ20から音が発生している間、図示せぬモータを駆動することによって、ぬいぐるみ1の嘴1aを開閉させる。これにより、ぬいぐるみ1はあたかもしゃべっているように見える。

20 【0053】次に、制御部29は、ステップS a 25へ進み、操作部3の決定ボタンが押下されたか否かを判断する。ここで、例えば使用者がスピーカ20から発せられた音を聞き、マイク10に話しかけた言葉と同じ言葉が発声された場合等に、操作部3の決定ボタンを押下すると、ステップS a 25の判断結果がYESとなって、ステップS a 26へ進む。また、所定時間決定ボタンが押下されなかった場合は、ステップS a 2へ戻り、上述した処理を再度繰り返す。使用者により決定ボタンが押下されて、制御部29がステップS a 26の処理へ進むと、制御部29は、音声データ変更部14に一時記憶されていた音声データを音声データ記憶部15のアドレスV(i)に記憶させ、学習モードの処理を終了する。

30 【0054】以上説明したように、学習モードにおいては、マイク10に入力された音声を30回に1回の割合でしか正確に発声しないため、入力された音声を単純にそのまま記憶するものよりも、インコらしさを演出でき、しかも意外性を持たせることができる。また、同じ音声が発声回数以上入力された場合は、正確に発声する割合が増すため、何度同じ音声を入力してもなかなか入力した通りに発声しないといった現象が減る。

40 【0055】さらには、ごく希に意図的に正確に発音する割合を増加させ、また、正確に発音する割合が増加したことを使用者に知らせることで、使用者にパチンコ遊技時におけるリーチ音のような感じで期待感を持たせることができる。また、マイク10から入力された音声がある時は、正確に発音する割合が増加するため、例えば当該玩具を持ち寄って、同じ玩具同士で互い音声を学習しあう場合は容易に音声を習得するといった区別化がなされ、使用者が自らの声を習得させる場合とはまた別の楽しみ方が可能となる。

50 【0056】〔設定モードにおける処理〕操作部3の設

定モードボタンが押下されると、制御部29は、図4および図5に示すフローチャートの処理を開始する。設定モードは、現在時刻、または、上述した学習モードで記憶させた音声をどのような場合に発声させるかの設定を行うモードであるが、その設定の内容には以下の5種類がある。

・設定1

使用者が発した特定の言葉または音に対し、記憶している特定の音声を発声させる。例えば、使用者が発した「ただいま」という言葉に対して、記憶している「お帰り」という言葉を発声させる。

【0057】・設定2

使用者が発した特定の言葉または音に対し、現在時刻を音声により通知させる。例えば、使用者が発した「何時?」という言葉に対して、「〇〇時△△分(現在時刻)だよ」という音声を発声させる。

・設定3

使用者が操作部3のテンキーにより設定した時刻に、現在時刻を音声により通知させる。

【0058】・設定4(目覚まし機能)

使用者が操作部3のテンキーにより設定した時刻に、予め指定された特定の音声を繰り返し発声させ、使用者の特定の言葉が入力されると、その音声の発声を停止する。例えば、指定された時刻になると「おはよう」という言葉を繰り返し発声し、使用者の「起きたよ」という言葉によって発声を停止する。

・設定5

使用者が操作部3のテンキーにより設定した時刻に、予め指定された特定の音声を発声させる。例えば、指定された時刻になると「テレビの時間だよ」という言葉を発声させる。

【0059】以下、上述した各種設定を行う時の制御部29の処理について、図4および図5を参照して説明する。まず、使用者が操作部3の設定モードボタンを押下すると、制御部29は、ステップSb1へ進み、マイク10から音声が入力されたか否かを判断する。そして、音声信号検出部22から検出信号が出力されない場合は、判断結果がNOとなり、ステップSb2へ進む。ステップSb2では、操作部3の時刻設定ボタンが押下されたか否かを判断する。ここで、操作部3の時刻設定ボタンが押下されなかった場合は、判断結果がNOとなり、ステップSb1の処理に戻る。このように、まず、制御部29は、音声が入力されるか、もしくは、時刻設定ボタンが押下されるまで、ステップSb1とステップSb2の処理を繰り返す。

【0060】この状態において、マイク10に音声が入力されると、音声入力検出部22から検出信号が出力され、これにより制御部29は音声が入力されたことを認識する。そして、音声入力検出部22から検出信号が出力されなくなった時点で、ステップSb3へ進む。そし

て、ステップSb3において、制御部29は操作部3のテンキーが押下されたか否かを判断し、テンキーが押下されなかった場合は判断結果がNOとなって、ステップSb4へ進む。ステップSb4では、操作部3の現在時刻ボタンが押下されたか否かを判断する。ここで、現在時刻ボタンが押下されなかった場合は、判断結果がNOとなり、ステップSb3の処理に戻る。このように、制御部29は、テンキーが押下されるか、もしくは、現在時刻ボタンが押下されるまで、ステップSb3とステップSb4の処理を繰り返す。

【0061】そして、例えば、テンキーが押下されたことを検知すると、制御部29はステップSb5の処理へ進み、設定1に関する処理を行う。まず、ステップSb5において、テンキーによる入力値をテンキーデータnとして記憶した後、ステップSb6へ進んで、音声データ記憶部15のアドレスV(n)に記憶されている音声データをD/A変換器16へ出力させる。これにより、テンキーによって指定されたアドレスに記憶された音声データがスピーカ20から発声される。

【0062】次にステップSb7へ進み、決定ボタンが押下されたか否かを判断する。ここで、決定ボタンが押下されなかった場合は、判断結果がNOとなり、ステップSb3へ戻る。そして、再度テンキーが押下された場合は、ステップSb5へ進み、テンキーにより指定されたアドレスに記憶されている音声データを、スピーカ20から発音させる。このようにして、使用者は、マイク10に入力した音声に対して発音させる音声を、テンキーによって選択する。

【0063】そして、所望する音声が発音された場合、使用者は、決定ボタンを押下する。すると、ステップSb7における判断結果がYESとなって、ステップSb8へ進み、マイク10に入力された音声と、その音声に対して本実施形態の玩具が発声する音声との対応が設定される。すなわち、ステップSb8において、制御部29は、選択されたアドレスV(n)を設定1を実行する際に読み出すアドレス、すなわち、読出アドレスRA1として記憶する。次いで、ステップSb9へ進み、特徴パターン抽出部21から出力された特徴パターンを、設定1を実行する際に、使用者が発声した音声を判断するための特徴パターンとして特徴パターン記憶部25のアドレスPS1に記憶させる。

【0064】その後、制御部29はステップSb10へ進み、使用者により設定1がなされたことを示す設定1フラグEST1を「1」にセットし、ステップSb11へ進む。ステップSb11では、使用者が設定モードにおいて、全ての設定が完了したか否かの判断を行う。すなわち、操作部3の決定ボタンが押下された場合は、判断結果がYESとなり、設定モードにおける処理を終了する。

【0065】また、決定ボタンが押下されなかった場合

は判断結果がNOとなって、ステップS b 1 2へ進み、決定ボタン以外のボタンが押下されたか否かの判断を行う。ここで、決定ボタン以外のボタンが押下された場合は判断結果がYESとなり、使用者が引き続き他の設定を行うものとみなして、ステップS b 1の処理に戻る。また、何等ボタンが押下されなかった場合は判断結果がNOとなり、ステップS b 1 1へ戻って、以下、決定ボタンまたはそれ以外のボタンが押下されるまで、ステップS b 1 1およびS b 1 2の処理を繰り返す。

【0066】次に、ステップS b 3においてテンキーが押下されず、かつ、ステップS b 4において操作部3の現在時刻ボタンが押下された場合は、ステップS b 4における判断結果がYESとなって、制御部29は設定2に関する処理を行う。すなわち、ステップS b 1 3へ進み、特徴パターン抽出部21から出力された特徴パターンを、設定2を実行する際、使用者が発声した音声来判断するための特徴パターンとして、特徴パターン記憶部25のアドレスPS2に記憶させる。

【0067】そして、ステップS b 1 4へ進み、使用者によって設定2がなされたことを示す設定2フラグEST2を「1」にセットし、ステップS b 1 1へ進む。そして、使用者が設定モードにおいて、全ての設定が完了したか否かの判断を行い、決定ボタンが押下された場合は、全ての設定が完了したとみなして、設定モードにおける処理を終了する。

【0068】さて、次にステップS b 1において音声信号が入力されないと判断され、かつ、ステップS b 2において操作部3の時刻設定ボタンが押下された場合は、ステップS b 2における判断結果がYESとなって、制御部29は設定3ないし5のいずれかに関する処理を行う。以下にその詳細について説明する。まず、ステップS b 2における判断結果がYESとなると、ステップS b 1 5へ進み、操作部3のテンキーから入力される値を、使用者によって指定された時刻とみなし、その値を指定時刻データTとして記憶する。

【0069】次にステップS b 1 6へ進み、操作部3の現在時刻ボタンが押下されたか否かを判断する。ここで、現在時刻ボタンが押下されると、判断結果はYESとなり、ステップS b 1 7へ進み、制御部29は設定3に関する処理を行う。すなわち、上述した指定時刻データTを、設定3における指定時刻T3として記憶する。そして、ステップS b 1 8へ進み、使用者によって設定3がなされたことを示す設定3フラグEST3を「1」にセットし、ステップS b 1 1へ進む。その後、上述したように、使用者が設定モードにおける全ての設定が完了したか否かの判断を行い、決定ボタンが押下された場合は、全ての設定が完了したとみなして、設定モードにおける処理を終了する。

【0070】また、ステップS b 1 6において、現在時刻ボタンが押下されなかった場合は、判断結果がNOと

なってステップS b 1 9へ進む。ステップS b 1 9では、マイク10から音声が入力されたか否かの判断を行う。ここで、音声入力検出部22から検出信号が出力された場合は、音声が入力されたとして判断結果がYESとなる。そして、ステップS b 2 0へ進み、以下、制御部29は設定4に関する処理を行う。まず、ステップS b 2 0において、制御部29は、特徴パターン抽出部21から出力された特徴パターンを、設定4を実行する際、使用者が発声した音声来判断するための特徴パターンとして、特徴パターン記憶部25のアドレスPS4に記憶させる。

【0071】次にステップS b 2 1へ進み、制御部29はテンキーの押下待機状態となる。すなわち、テンキーのいずれかが押下されるまでは判断結果がNOとなって、ステップS b 2 1の処理を繰り返し、テンキーが押下されると判断結果がYESとなって、ステップS b 2 2へ進む。そして、ステップS b 2 2では、テンキーの入力値をテンキーデータnに代入した後、ステップS b 2 3へ進んで、音声データ記憶部15のアドレスV(n)に記憶されている音声データをD/A変換器16へ出力させる。これにより、テンキーによって指定されたアドレスに記憶された音声データが、スピーカ20から発声される。

【0072】次にステップS b 2 4へ進み、決定ボタンが押下されたか否かを判断する。ここで、決定ボタンが押下されなかった場合は、判断結果がNOとなり、ステップS b 2 1へ戻って、再度テンキーの押下待機状態となる。以下、制御部29は、決定ボタンが押下されるまで、ステップS b 2 1～S b 2 4の処理を繰り返す。これにより、使用者は、テンキーによって設定4を実行する際に発声させる所望の音声を、テンキーによって選択することができる。

【0073】そして、スピーカ20から使用者が所望する音声が発せられ、使用者が決定ボタンを押下すると、ステップS b 2 4における判断結果がYESとなって、制御部29はステップS b 2 5以降の処理を行う。まず、ステップS b 2 5において、制御部29は、上述したアドレスV(n)を設定4を実行する際に読み出すアドレス、すなわち、読出アドレスRA4として記憶する。次いで、ステップS b 2 6へ進み、ステップS b 1 5において使用者により指定された時刻データTを、設定4における指定時刻T4として記憶する。

【0074】その後、制御部29はステップS b 2 7へ進み、使用者により設定4がなされたことを示す設定4フラグEST4を「1」にセットし、ステップS b 1 1へ進む。ステップS b 1 1では、使用者が設定モードにおいて、全ての設定が完了したか否かの判断を行う。すなわち、操作部3の決定ボタンが押下された場合は、判断結果がYESとなり、使用者が全ての設定を完了したとみなして設定モードにおける処理を終了する。

10

20

30

40

50

【0075】次に、ステップS b 1 9において、マイク10に音声が入力されなかったと判断された場合、制御部29はステップS b 2 8へ進み、テンキーが押下されたか否かを判断する。そして、テンキーが押下された場合は判断結果がYESとなって、ステップS b 2 9以降の設定5に関する処理を行う。なお、ステップS b 2 8において、テンキーが押下されなかった場合は、判断結果がNOとなって、ステップS b 1 6へ戻る。これにより、制御部29は、操作部3の設定モードボタンが押下され、まず最初に、時刻設定ボタンが押下されてテンキーより時刻が設定された後は、現在時刻ボタンの押下、本実施形態の玩具に対する音声の入力、または、テンキーの押下のいずれかに応じて、設定3ないし5のいずれかの処理を行うことになる。

【0076】制御部29がステップS b 2 9の処理へ進むと、制御部29は、テンキーの入力値を変数nに代入した後、ステップS b 3 0へ進んで、音声データ記憶部15のアドレスV(n)に記憶されている音声データをD/A変換器16へ出力させる。これにより、テンキーによって指定されたアドレスに記憶された音声データがスピーカ20から発声される。次に制御部29はステップS b 3 1へ進み、決定ボタンが押下されたか否かを判断する。ここで、決定ボタンが押下されなかった場合は、判断結果がNOとなり、ステップS b 1 6へ戻る。また、ステップS b 1 6→S b 1 9→S b 2 8の処理を経る間にテンキーが押下されると、再びステップS b 2 8の判断結果がYESとなって、ステップS b 2 9の処理が行われる。

【0077】そして、スピーカ20から、使用者が所望する音声が発せられて、使用者が決定ボタンを押下すると、ステップS b 3 1における判断結果がYESとなって、制御部29はステップS b 3 2以降の処理を行う。ここで、ステップS b 3 2以降の処理では、使用者によって指定された時刻と、その時刻において本実施形態の玩具が発声する音声との対応が設定される。まず、ステップS b 3 2において、制御部29は、上述したアドレスV(n)を設定5を実行する際に読み出すアドレス、すなわち、読出アドレスRA5として記憶する。次いで、ステップS b 3 3へ進み、ステップS b 1 5において使用者により指定された時刻データTを、設定5における指定時刻T5として記憶する。

【0078】その後、制御部29はステップS b 3 4へ進み、使用者により設定5がなされたことを示す設定5フラグEST5を「1」にセットし、ステップS b 1 1へ進む。ステップS b 1 1では、使用者が設定モードにおいて、全ての設定が完了したか否かの判断を行う。すなわち、操作部3の決定ボタンが押下された場合は、判断結果がYESとなり、使用者が全ての設定を完了したとみなして設定モードにおける処理を終了する。

【0079】以上説明したように、設定モードにおいて

は、操作部3の設定モードボタンが押下された後、まず最初にマイク10に音声が入力され、引き続きテンキーが押下された場合は設定1の処理が行われ、現在時刻ボタンが押下された場合は設定2の処理が行われる。また、設定モードボタンが押下されて、まず最初に時刻設定ボタンが押下され、次いで時刻設定がなされた場合、引き続き、現在時刻ボタンが押下された場合は設定3の処理を、音声が入力された場合は設定4の処理を、また、テンキーが押下された場合は設定5の処理を行う。

10 【0080】〔実行モードにおける処理〕操作部3の実行モードボタンが押下されると、制御部29は、図6および図7に示すフローチャートの処理を開始する。実行モードは上述した設定モードにおいて設定された内容を実行するモードであり、以下、実行モードにおける制御部29の動作について説明する。実行モードボタンが押下されると、制御部29は、まず、図6(a)のフローチャートにおけるステップSc1の処理を行う。すなわち、各設定フラグEST1～EST5の状態をサーチし、「1」がセットされているフラグを検出する。

20 【0081】そして、ステップSc2へ進み、タイマ23に対し現在時刻の情報を出力させ、その情報を現在時刻データTcとして記憶する。そして、ステップSc3へ進み、ステップSc1でサーチした結果に基づいて、「1」がセットされている設定フラグについて、その設定内容をそれぞれ1回ずつ実行する。なお、ステップSc3における処理については後述する。次にステップSc4に進み、操作部3のランダムモードボタンが押下されているか否かを判断する。

30 【0082】そして、ランダムモードボタンが押下されていた場合は、判断結果がYESとなって、ステップSc5に進み、ランダムモードの処理を実行する。なお、この処理については、後に詳しく説明する。また、ランダムモードボタンが押下されていなかった場合は、判断結果がNOになって、ステップSc5の処理を行わずにステップSc6へ進み、他のモードボタンが押下されたか否かを判断する。ここで、他のモードボタンが押下された場合は、判断結果がYESとなって、実行モードの処理を終了すると共に、押下されたモードボタンのモードを実行する。また、ステップSc6において、他のモードボタンが押下されなかった場合は、判断結果がNOとなって、ステップSc2へ戻り、以下、ステップSc2～Sc6の処理を繰り返し実行する。

【0083】次に、上述したステップSc3において、各設定を実行する際の処理動作について図6および図7を参照して説明する。まず、設定1フラグEST1が「1」にセットされていた場合、制御部29は、図6

(b)に示すフローチャートの処理を実行する。ステップSc7において、制御部29は、マイク10に音声が入力されたか否かの判断を行う。ここで、音声入力検出部22から検出信号が出力され、音声が入力されたと判

断すると、判断結果はYESとなり、ステップSc 8の処理を行う。また、音声入力検出部22から検出信号が出力されなかった場合は、判断結果がNOとなり、図6(a)のステップSc 3へ戻り、「1」にセットされている他の設定フラグに関する実行処理を行う。

【0084】そして、ステップSc 7における判断結果がYESとなって、ステップSc 8へ進むと、制御部29は特徴パターン抽出部21から出力された特徴パターンと、特徴パターン記憶部25のアドレスPS 1に記憶されている特徴パターンと比較する。そして、ステップSc 9へ進み、両者の特徴パターンが一致したか否かを判断する。両者の特徴パターンが一致しなかった場合は、判断結果がNOとなり、図6(a)のステップSc 3へ戻って、「1」にセットされている他の設定フラグに関する実行処理を行う。また、一致した場合は、判断結果がYESとなってステップSc 10へ進み、読出アドレスRA 1によって特定される、音声データ記憶部15のアドレスに記憶された音声データを、D/A変換器16へ出力するよう指示する。

【0085】そして、D/A変換器16へ出力された音声データは、アナログ音声信号に変換され、音声変換回路17においてインコのような音色に変換された後、増幅回路18および音量調整回路19を経て、スピーカ20から音声が発せられる。これにより、例えば、特徴パターン記憶部25のアドレスPS 1に「ただいま」という音声に相当する特徴パターンが記憶され、読出アドレスRA 1によって特定される音声データ記憶部15のアドレスに「お帰り」という音声データが記憶されていた場合、使用者がマイク10に「ただいま」という音声を入力すると、スピーカ20から「お帰り」という音声が発せられる。

【0086】また、この時、制御部29は駆動装置27に対してぬいぐるみ1の嘴1aを駆動するよう指示し、これにより、スピーカ20から音が発生している間、ぬいぐるみ1の嘴1aが開閉し、ぬいぐるみ1があたかもしゃべっているように見える。そして、ステップSc 10の処理を終えると、制御部29は、図6(a)のステップSc 3へ戻り、「1」にセットされている他の設定フラグに関して、その設定の実行処理を行う。

【0087】次に、設定2フラグEST 2が「1」にセットされていた場合における、制御部29の処理動作について、図6(c)に示すフローチャートを参照して説明する。まず、制御部29は、ステップSc 11へ進み、マイク10に音声が入力されたか否かの判断を行う。ここで、音声入力検出部22から検出信号が出力され、音声が入力されたと判断すると、判断結果はYESとなり、ステップSc 12の処理を行う。また、音声入力検出部22から検出信号が出力されなかった場合は、判断結果がNOとなり、図6(a)のステップSc 3へ戻り、「1」にセットされている他の設定フラグにつ

て、その設定の実行処理を行う。

【0088】そして、ステップSc 11における判断結果がYESとなって、ステップSc 12へ進むと、制御部29は特徴パターン抽出部21から出力された特徴パターンと、特徴パターン記憶部25のアドレスPS 2に記憶されている特徴パターンと比較する。そして、ステップSc 13へ進み、両者の特徴パターンが一致したか否かを判断する。両者の特徴パターンが一致しなかった場合は、判断結果がNOとなり、図6(a)のステップSc 3へ戻って、「1」にセットされている他の設定フラグについて、その設定の実行処理を行う。また、一致した場合は判断結果がYESとなって、ステップSc 14へ進み、音声データ記憶部15に対し、現在時刻データTcに対応する音声データをD/A変換器16へ出力するよう指示する。

【0089】そして、D/A変換器16へ出力された現在時刻に対応する音声データは、アナログ音声信号に変換され、音声変換回路17においてインコのような音色に変換された後、増幅回路18および音量調整回路19を経て、スピーカ20から現在時刻が発せられる。これにより、例えば、特徴パターン記憶部25のアドレスPS 2に「何時?」という音声に相当する特徴パターンが記憶されていた場合、使用者がマイク10に「何時?」という音声を入力すると、スピーカ20からは、現在時刻がインコのような音色で発せられる。

【0090】そして、ステップSc 14の処理を終えると、制御部29は、図6(a)のステップSc 3へ戻り、「1」にセットされている他の設定フラグについて、その設定の実行処理を行う。

【0091】次に、設定3フラグEST 3が「1」にセットされていた場合における、制御部29の処理動作について、図7(a)に示すフローチャートを参照して説明する。まず、制御部29は、ステップSc 15へ進み、設定モードにおける設定3で使用者によって設定された指定時刻データT3と、現在時刻データTcとを比較し、一致するか否かを判断する。そして、両者が一致した場合は、判断結果がYESとなり、ステップSc 16へ進み、音声データ記憶部15に対し、現在時刻データTcに対応する音声データを、D/A変換器16へ出力するよう指示する。

【0092】そして、D/A変換器16へ出力された、現在時刻に対応する音声データは、アナログ音声信号に変換され、音声変換回路17においてインコのような音色に変換された後、増幅回路18および音量調整回路19を経て、スピーカ20から現在時刻が発せられる。これにより、本実施形態における玩具は、設定モードの設定3で指定された時刻になると、現在時刻をインコのような声で発声する。そして、ステップSc 16の処理を終えると、制御部29は、図6(a)のステップSc 3へ戻り、「1」にセットされている他の設定フラグにつ

いて、その設定の実行処理を行う。一方、ステップSc 15において、指定時刻データT 3と現在時刻データT cが一致しなかった場合は、判断結果がNOとなって、ステップSc 16の処理を行わずに、図6 (a)のステップSc 3へ戻る。

【0093】次に、設定4フラグEST 4が「1」にセットされていた場合における、制御部29の処理動作について、図7 (b)に示すフローチャートを参照して説明する。まず、制御部29は、ステップSc 17へ進み、操作部3の目覚ましボタンが押下されているか否かの判断を行う。そして、目覚ましボタンが押下されてい

なかった場合は、判断結果がNOとなり、図6 (a)のステップSc 3へ戻り、「1」にセットされている他の設定フラグについて、その設定の実行処理を行う。

【0094】また、目覚ましボタンが押下されていた場合は、判断結果がYESとなり、ステップSc 18の処理へ進む。そして、ステップSc 18において、制御部29は、設定モードにおける設定4で、使用者によって設定された指定時刻データT 4と、現在時刻データT cとを比較し、一致するか否かを判断する。そして、両者が一致した場合は、判断結果がYESとなってステップSc 19へ進み、また、両者が一致しなかった場合は、判断結果がNOとなって図6 (a)のステップSc 3へ戻る。

【0095】ステップSc 19へ進むと、制御部29は、読出アドレスRA 4によって特定される、音声データ記憶部15のアドレスに記憶された音声データを、D/A変換器16へ出力するよう指示する。そして、D/A変換器16へ出力された音声データは、アナログ音声信号に変換され、音声変換回路17においてインコのような音色に変換された後、増幅回路18および音量調整回路19を経て、スピーカ20から音声が発せられる。

【0096】次いで、制御部29はステップSc 20へ進み、マイク10に音声が入力されたか否かを判断する。ここで、音声入力検出部22から検出信号が出力され、音声が入力されたと判断すると、判断結果はYESとなり、ステップSc 21の処理を行う。また、音声入力検出部22から検出信号が出力されなかった場合は、判断結果がNOとなり、再度ステップSc 19へ戻って、読出アドレスRA 4によって特定される、音声データ記憶部15のアドレスに記憶された音声データを発声させるための処理を行う。

【0097】そして、ステップSc 20における判断結果がYESとなって、ステップSc 22へ進むと、制御部29は特徴パターン抽出部21から出力された特徴パターンと、特徴パターン記憶部25のアドレスPS 4に記憶されている特徴パターンと比較する。そして、ステップSc 22へ進み、両者の特徴パターンが一致したか否かを判断する。両者の特徴パターンが一致しなかった場合は、判断結果がNOとなり、ステップSc 19へ戻

って、再度読出アドレスRA 4によって特定される、音声データ記憶部15のアドレスに記憶された音声データに基づいて発声させるための処理を行う。

【0098】また、特徴パターン抽出部21から出力された特徴パターンと、特徴パターン記憶部25のアドレスPS 4に記憶されている特徴パターンとが一致すると、判断結果がYESとなって図6 (a)のステップSc 3へ戻る。このように、図7 (b)のフローチャートにおいては、操作部3の目覚ましボタンが押下されており、かつ、設定モードの設定4で指定された時刻になると、使用者が所定の音声を発声するまで、予め指定された音声スピーカ20から繰り返し発声する。

【0099】次に、設定5フラグEST 5が「1」にセットされていた場合における、制御部29の処理動作について、図7 (c)に示すフローチャートを参照して説明する。まず、制御部29は、ステップSc 23において、設定モードにおける設定5で使用者によって設定された指定時刻データT 5と、現在時刻データT cとを比較し、一致するか否かを判断する。そして、両者が一致した場合は、判断結果がYESとなってステップSc 24へ進み、また、両者が一致しなかった場合は、判断結果がNOとなって図6 (a)のステップSc 3へ戻る。

【0100】ステップSc 24へ進むと、制御部29は、読出アドレスRA 5によって特定される、音声データ記憶部15のアドレスに記憶された音声データを、D/A変換器16へ出力するよう指示する。そして、D/A変換器16へ出力された音声データは、アナログ音声信号に変換され、音声変換回路17においてインコのような音色に変換された後、増幅回路18および音量調整回路19を経て、スピーカ20から音声が発せられる。その後、制御部29は図6 (a)のステップSc 3へ戻る。このように、図7 (c)のフローチャートにおいては、設定モードの設定5で指定された時刻になると、使用者によって指定された音声が発声する。

【0101】なお、上述した実行モードにおいて、設定モード時に対応づけられなかった音声が入力されると、本実施形態における玩具は無反応となるが、そのような場合、例えば、制御部29が音声データ記憶部15に記憶された音声データの中から任意の音声データを選択し、D/A変換器16へ出力させるようにしてもよい。

【0102】〔ランダムモードにおける処理〕ランダムモードの処理は、図6 (a)のフローチャートからも明かなように、実行モードの実施中になされる。以下、図8のフローチャートを参照してランダムモードにおける処理について説明する。まず、図6 (a)のフローチャートのステップSc 4において、操作部3のランダムモードボタンが押下されていると判断された場合、制御部29は、ステップSc 5へ進み、ランダムモードにおける処理を開始する。すなわち、まず、図8のフローチャートにおいて、ステップSd 1へ進み、マイク10に対

10

20

30

40

50

して音声が入力されたか否かを判断する。

【0103】ここで、音声入力検出部22から検出信号が出力されて、マイク10に対して音声が入力されたと判断した場合は、判断結果がYESとなって入力された音声を記憶するか否かが乱数に基づいて決定される。すなわち、ステップSd2へ進み、乱数発生部27に乱数RNDを出力させる。次に、ステップSd3へ進み、乱数発生部27から出力された乱数RNDと定数dとを比較する。ここで、定数dは、0～1の数値範囲内で小数点第3位まで表示された所定の小数である。そして、乱数RNDと定数dとが一致した場合は、判断結果がYESとなってステップSd4へ進み、入力された音声の音声データを、音声データ記憶部15のアドレスVR

(j)に記憶させた後、ステップSd5へ進み、jの値を1インクリメントする。ここで、上述したアドレスVRは、音声データ記憶部15において、ランダムモード時に記憶した音声データが記憶される、メモリ領域を意味している。

【0104】そして、ステップSd1にて音声信号が入力されなかったと判断された場合(判断結果がNO)、ステップSd3にて乱数RNDと定数dとが一致しなかった場合(判断結果がNO)、もしくは、ステップSd5の処理が終了した場合は、ステップSd6へ進み、音声データ記憶部15に記憶されているいずれかの音声データを発生させるか否かを、乱数に基づいて決定する。すなわち、ステップSd6において乱数発生部27に乱数RNDを出力させ、次いで、ステップSd7で乱数発生部27から出力された乱数RNDが、定数eと一致するか否かを判断する。ここで、定数eは、0～1の数値範囲内で小数点第3位まで表示された、定数dとは異なる所定の小数である。

【0105】そして、乱数RNDと定数eが一致しなかった場合は、判断結果がNOとなり、図6(a)のフローチャートにおけるステップSc6へ戻る。また、乱数RNDと定数eが一致した場合は判断結果がYESとなり、ステップSd8へ進み、音声データ記憶部15に記憶された音声データのうち、どの音声データを発声するかを、乱数に基づいて決定する。すなわち、まず、乱数発生部27に乱数RNDを出力させ、この乱数RNDとステップSd5で示した変数jとに基づいて、0～jの数値範囲内の乱数を算出する。

【0106】次にステップSd9へ進み、音声データ記憶部15に対し、ステップSd8で得た乱数によって特定されるアドレスVR(k)から音声データを出力させる。これにより、音声データ記憶部15から出力された音声データは、D/A変換器16によってアナログ音声信号に変換され、音声変換回路17においてインコのような音色に変換された後、増幅回路18および音量調整回路19を経て、スピーカ20から音声が発せられる。そして、ステップSd9の処理を終えると、制御部29

は図6(a)のステップSc6へ進み、実行モードにおける処理を継続する。

【0107】このように、使用者によってランダムモードが選択された場合は、図6(a)のフローチャートのステップSc5を通過する毎に、図8に示すフローチャートの処理がなされる。これにより、本実施形態における玩具は、マイク10に入力された音声を任意に取り込み、また、任意の言葉を任意に発声する。

【0108】〔メッセージモードにおける処理〕操作部3のメッセージモードボタンが押下されると、制御部29は、図9に示すフローチャートの処理を開始する。まず、制御部29はステップSe1において、メッセージモードにおける音声データ記憶部15のアドレス指定用変数mの値をクリアする。そして、ステップSe2へ進み、音声データ記憶部15に対し、アドレスV(m)に記憶された音声データをD/A変換器16へ出力させる。これにより、D/A変換器16へ出力された音声データはアナログ音声信号に変換され、音声変換回路17においてインコのような音色に変換された後、増幅回路18および音量調整回路19を経て、スピーカ20から音声が発せられる。

【0109】次にステップSe3へ進み、アドレス指定用変数mの値を1インクリメントし、ステップSe4において、mの値が学習モードにおけるアドレス指定用変数iを超えたか否かを判断する。そして、mの値が変数iの値以下であった場合は、判断結果がNOとなり、ステップSe2へ戻って、音声データ記憶部15の、次のアドレスに記憶された音声データをD/A変換器16へ出力させる。このように、mの値を1つずつ増加させて学習モード時に本実施形態の玩具に習得させた音声データを順次スピーカ20から発声させる。そして、ステップSe4において、mの値が変数iの値を超えたと判断した場合は、判断結果がYESとなってメッセージモードにおける処理を終了する。

【0110】このように、メッセージモードにおいては、学習モード時に音声データ記憶部15に記憶させた音声データを順々に全て発声させるので、本実施形態の玩具が記憶している音声データの内容を確認できる。また、例えば、相手に伝えたいメッセージを学習モード時において本実施形態の玩具に習得させておけば、その後、その相手がメッセージモードボタンを押下した時に、習得させたメッセージを聞くことができるので、本実施形態の玩具をインコのような音色に変えてのメッセージャーとしても利用することができる。

【0111】以上説明したように、本実施形態における玩具によれば、入力された音声および音に対し、応答する音声を任意に設定することができる。これにより、玩具の反応に意外性を持たせることができ、飽きにくい玩具を提供することができる。

【0112】なお、上述した実施形態においては、イン



コを模したぬいぐるみを用いたが、インコに限らず、おうむ、九官鳥等の人語を発する動物の他、犬や猫等の他の動物のぬいぐるみや、人形等を用いてもよく、また、それら人形またはぬいぐるみの容積に余裕がある場合は、それらの内部に上述した音声認識および音声発生装置を組み込んでよい。また、上述した音声認識および音声発生装置を、室内調度品等に組み込んでよい。

【0113】また、音声変換回路17によってD/A変換器16から出力されたアナログ信号の波形を調節し、音声データ記憶部15から読み出された音声データをインコのような声で発生するようにしているが、音声変換回路17において、複数種類の音質に変換し得るように構成し、選択スイッチ等を設けて変換する音質を選択できるようにしてもよい。

【0114】さらに、上述した玩具の機能をパーソナルコンピュータ（以下、パソコンという）またはテレビゲーム機上で実現してもよい。その場合、ぬいぐるみ1の代わりとして、モニタ画面に、例えばインコ等の画像を表示させ、パソコンから音声を発生させる時に、画面上のインコの嘴や羽等を動かすような画像処理を行うようにしてもよい。ここで、図10に、パソコンを用いて上述した玩具の機能を実現する場合の構成を示す。

【0115】図10において、10はマイクロホン、20はスピーカ、50はパソコン本体、51はキーボード、52はモニタである。また、53はパソコン本体50に組み込まれる音源ボードであり、増幅回路、フィルタ、A/D変換器およびD/A変換器等を具備し、マイクロホン10から出力される音声信号をデジタル化してパソコン本体50内のバスに出力すると共に、バスからのデジタルデータをアナログ化してスピーカ20へ出力する。

【0116】図10に示す構成においては、図2に示す増幅回路11、18、A/D変換器13、D/A変換器16は、音源ボード53に対応し、図2の操作部3は、キーボード52またはマウス（図示略）等の入力デバイスに対応する。また、図2における音声データ記憶部15、玩具特徴パターン記憶部24、特徴パターン記憶部25は、パソコン本体50内のRAMを利用することで実現可能である。さらに、図2の音声データ変更部14、特徴パターン変換部21、音声入力検出部22における処理は、音源ボード53から出力されるデジタル化された音声信号に基づいて、パソコン本体50により行われる。また、図2のタイマ23、カウンタ26、乱数発生部27の機能は、パソコン本体50が具備する内蔵時計、および、RAM等を利用することによって実現可能である。

【0117】そして、図3ないし図9のフローチャートに示した処理手順をパソコン本体50に実行させるためのプログラムを、例えば、フロッピーディスク等の記録媒体に記憶させておき、その記録媒体から上記プログラ

ムを読み出し、実行することで、前述した玩具の機能を図10に示す構成により実現させる。なお、パソコンを用いて前述した玩具の機能を実現する場合は、既に述べたように、図1に示すぬいぐるみ1の代わりとして、モニタ51に例えばインコの画像を表示させ、スピーカ20から音声を発生させる時に画面上のインコの嘴や羽等を動かすような画像処理を行うので、図3のステップSa24に示した駆動装置を駆動する処理の代わりに上述した画像処理を行うことになる。

【0118】〔他の実施形態〕以下に、上述した各モードのうち、学習モードと、設定モードにおける設定1（使用者が発した特定の言葉または音に対し、記憶している特定の音声を発声させるための設定）の他の実施形態について説明する。

【0119】〔学習モードの他の実施形態〕図11および図12を参照して、図3に示した学習モードの他の形態について説明する。この実施形態における学習モードでは、玩具が1つの言葉を学習するにあたり、学習の開始から終了するまでの間に複数の段階（学習段階）を設け、現在学習中の言葉がどの段階にあるのかを、ユーザが判別できるようにしたものである。なお、本実施形態においては、少なくとも当該玩具につけられた名前に対応する言葉の特徴パターンが、予め特徴パターン記憶部25に記憶されているものとする。

【0120】まず、本実施形態の玩具の電源スイッチがONされ、制御部29が操作部3のモード選択ボタンの押下待機状態になった時に学習モードボタンが押下されると、制御部29は、学習モードの処理を開始する。

【0121】まず、図11のステップSf1へ進み、音声登録処理を開始し、次いでステップSf2で、マイクロホン10から音声入力があるか否かを判断する。ここで、音声が入力されなかった場合は、判断結果がNoとなり、再度ステップSf2の処理を行い、以下、言葉（または音）が入力されるまで音声入力待機状態となる。そして、マイクロホン10に音声が入力され、音声入力検出部22から検出信号が検出されると、判断結果がYesとなってステップSf3へ進み、A/D変換器13から出力された音声データが音声データ変更部14に一時記憶される。

【0122】次に、ステップSf4に進んで、乱数を発生し、ステップSf5で、発生した乱数の値と所定値 $\alpha$ とを比較する。もし、乱数の値が所定値 $\alpha$ 以上であった場合は、判断結果がNoとなり、ステップSf1へ戻る。また、乱数の値が所定値 $\alpha$ よりも小さかった場合は、判断結果がYesとなってステップSf6へ進み、これ以降入力される音声の学習を行う。

【0123】ここで、上述したステップSf1～Sf5の処理は、本実施形態における玩具に、より本物らしいインコを模擬させるため、学習モードボタンが押下されると必ず言葉（または音）の学習を開始するのではな

く、乱数に基づいて、学習を開始しない場合を生じさせるために設けられている。

【0124】ステップS f 6に移行すると、制御部29は、ステップS f 2で入力された音声の特徴パターンを、特徴パターン記憶部25へ登録する。また、これと共に、音声データ変更部14に一時記憶されていた音声データを“学習中の言葉”として音声データ記憶部15に記憶する。次にステップS f 7へ進み、“学習の言葉”として記憶された音声データに対応する、段階パラメータSTEP、入力回数レジスタIN、記憶保持パラメータMHをそれぞれ発生し、各パラメータおよびレジスタの値をそれぞれ0にセットする。

【0125】ここで、段階パラメータSTEPは、“学習中の言葉”の学習段階を示すパラメータであり、0～3の整数の値をとり、0～2までが学習中の言葉であることを意味し、3は学習済みの言葉であることを意味する。また、このパラメータの値に応じたぬいぐるみ1の動作が予め設定されている。すなわち、例えば、段階パラメータSTEPの値が、0の時は嘴を1回開閉させる、1の時は翼を羽ばたかせる、2の時は嘴を繰り返し開閉させると共に翼を羽ばたかせる、3の時は2の時の動作に加え、インコの鳴き声を発生させる等の動作が予め設定されている。

【0126】また、入力回数レジスタINは、ステップS f 6以降に入力された言葉の入力回数が記憶されるレジスタである。さらに、記憶保持パラメータMHは、音声データ記憶部に記憶された音声データを消去するか否かを決定するためのパラメータである。なお、記憶保持パラメータMHの詳細については後述する。

【0127】そして、上述したパラメータおよびレジスタを発生した後、段階パラメータSTEPの値に応じた動作をぬいぐるみ1に実行させる。そして、ステップS f 8へ進み、音声認識を開始する。すなわち、ステップS f 9へ進んで、ステップS f 2と同様に音声入力待機状態となり、マイク10に音声が入力されると、ステップS f 10へ進み、特徴パターン抽出部21から出力された、その音声の特徴パターンと、特徴パターン記憶部25に登録されている各特徴パターンとを順次比較し、入力された音声の認識を行う。

【0128】次に、ステップS f 11で乱数を発生し、ステップS f 12で、発生した乱数の値が、前述した所定値 $\alpha$ よりも小さいか否かが判断される。もし、乱数RNDの値が所定値 $\alpha$ 以上であれば、判断結果がNoとなってステップS f 8へ戻る。また、もし、乱数の値が所定値 $\alpha$ 未満であれば、判断結果がYesとなり、図12のステップS f 13へ進む。

【0129】そして、ステップS f 13において、ステップS f 10で認識した音声が、本実施形態の玩具に付けられた名前であるか否かを判断する。ここで、もし玩具につけられた名前であった場合は、判断結果がYes

となり、本学習モードが終了する。また、玩具につけられた名前でなかった場合は、判断結果がNoとなり、ステップS f 14へ進んで、“学習中の言葉”（段階パラメータSTEPの値が2以下の言葉）であるか否かを判断し、“学習中の言葉”でなかったと判断された場合は、判断結果がNoとなり、認識された言葉に応じた動作をぬいぐるみ1に実行させ（ステップS f 15）した後、ステップS f 8へ戻る。

【0130】ここで、ステップS f 14で“学習中の言葉”でないと判断される場合は、認識された言葉（または音）が、学習済みの言葉（段階パラメータSTEPの値が3言葉）であるか、“学習中の言葉”以外の言葉（段階パラメータSTEPが与えられていない言葉）であるかの、いずれかの場合であり、制御部29は、ステップS f 15において、各々の場合について予め定められている動作をぬいぐるみ1に実行させる。

【0131】一方、ステップS f 14で、“学習中の言葉”であると判断された場合は、判断結果がYesとなって、ステップS f 16へ進み、乱数を発生する。そしてステップS f 17へ進み、乱数の値が所定値 $\beta$ 未満か否かを判断する。ここで、所定値 $\beta$ は、学習段階を1段階進めるか否かを決定するために設けられている数値である。

【0132】そして、乱数の値が所定値 $\beta$ 未満であった場合は、判断結果がYesとなって、ステップS f 18へ進み、段階パラメータSTEPの値を1つインクリメントすると共に、記憶保持パラメータMHの値に所定の固定値（例えば10）を加算する。また、入力回数レジスタINの値を0にリセットする。この処理によって、音声データ記憶部15に、“学習中の言葉”として記憶された音声データの学習段階が、1段階進められたことになる。

【0133】また、前述したステップS f 17において、乱数の値が所定値 $\beta$ 以上であった場合は、判断結果がNoとなってステップS f 20へ移行する。そして、入力回数レジスタINの値を1つインクリメントして、ステップS f 21で入力回数レジスタINの値が所定値 $\gamma$ よりも大きくなったか否かの判断を行う。もし、入力回数レジスタINの値が所定値 $\gamma$ よりも大きければ、判断結果がYesとなってステップS f 18へ進み、段階パラメータSTEPの値を1インクリメントして、認識された“学習中の言葉”の学習段階を1段階進める。このように、ステップS f 21の処理により、同一の“学習中の言葉”が、ある回数以上認識された場合は、強制的に学習段階が1段階進められる。

【0134】また、ステップS f 18で学習段階が1段階進められた後は、ステップS f 19へ進んで、段階パラメータSTEPの値が3になったか否かの判断を行う。段階パラメータSTEPの値が2以下であれば、判断結果がNoとなり、認識された“学習中の言葉”の学

習が終了していないものとみなしてステップS f 8へ戻る。

【0135】一方、ステップS f 21において、入力回数レジスタINの値が所定値 $\gamma$ 以下であれば、判結果がNoとなって、ステップS f 22で、段階パラメータSTEPの値に応じた動作をぬいぐるみ1に実行させる。これによりユーザは、ステップS f 10で認識された音声、現在どの段階にあるのかを認識することができる。そして、ステップS f 22の処理が終了すると、図11のステップS f 8へ戻る。

【0136】以後、ステップS f 1～S f 22の処理を繰り返し行い、ステップS f 18で段階パラメータSTEPの値が3になると、ステップS f 19の判断結果がYesとなって、ステップS f 23へ進み、段階パラメータSTEPの値(=3)に対応した動作を、ぬいぐるみ1に実行させる。これにより、ユーザは、1つの“学習中の言葉”が学習されたことを認識することができる。そして、ステップS f 24へ進み、記憶保持パラメータMHの値を40にセットして学習モードを終了する。

【0137】なお、上述した記憶保持パラメータMHの値を、他のモード、例えば実行モード等で、その記憶保持パラメータMHに対応する音声認識された時に、1つインクリメント(ただし1日に1回のみ)し、また、1日経過する毎に1つデクリメントして、値が0になった時に、その音声データおよび特徴パターンをそれぞれの記憶部から消去する。これにより、本実施形態における玩具に、入力頻度の低い言葉または音については、自然に忘れてしまうように振る舞わせることができる。

【0138】このように、本実施形態の学習モードにおいては、実質的な学習モードの処理(すなわち、ステップS f 6以降の処理)に移行する前に、言葉または音の記憶(正確には、音声データおよび特徴パターンの記憶)が1度だけ行われ、それ以降に入力される言葉によって、記憶された言葉または音(過去に記憶されたものも含む)の学習が行われる。また、言葉の学習は、玩具につけられた名前が入力されるか、または、1つの学習が終了するまで継続して行われ、さらに、学習段階が1段階進む毎にそれに応じた動作を行うので、ユーザは、学習の進み具合を知ることができる。

【0139】〔設定モード〔設定1〕の他の実施形態〕次に、設定モード〔設定1〕の他の実施形態について、図13および図14を参照して説明する。先に説明した設定モード〔設定1〕の処理(図4、ステップS b 1、S b 3、S b 5～S b 11)では、まず、ユーザが音声を入力し、この入力した音声に対応付けたい音声を、ユーザ自身が選択して決定キーを押下することで音声の対応付けが行われていたが、本実施形態では、このような操作を必要とせずに、音声の対応付けを可能にしたものである。

【0140】まず、設定モード〔設定1〕の処理が開始されると、図13のステップS g 1において乱数を発生し、次にステップS g 2に進み、音声データ記憶部15の乱数RND番目のアドレスに、学習済みの音声データ(段階パラメータSTEPの値が3の音声データ)が記憶されているか否かを判断する。そして、記憶されていなかった場合は、判断結果がNoとなってステップS g 1へ戻り、学習済みの音声データが記憶されているアドレスが見つかるまで、ステップS g 1およびS g 2の処理を繰り返す。

【0141】一方、ステップS g 2で、学習済みの音声データを記憶したアドレスが見つかったと、判断結果がYesとなってステップS g 3へ進む。ここで、その音声データに対応する記憶保持パラメータMHの値を読み取り、ステップS g 4で、記憶保持パラメータMHの値が40よりも大きいのか否かの判断が行われる。ここで、40以下と判断された場合は、判断結果がNoとなって、ステップS g 1へ戻る。また、40よりも大きいと判断された場合は、ステップS g 5へ進み、当該アドレスに記憶されていた音声データをD/A変換器16へ出力し、スピーカ20から発声され、この時、制御部29は、所定時間(ここでは7秒間とする)の計時を開始する。

【0142】次に、ステップS g 6へ進み、制御部29は、これ以降に入力される音声の認識を行う。すなわち、ステップS g 7でマイク10に音声が入力されたか否かの判断を行い、音声の入力がない場合は、ステップS g 8において、計時開始後、7秒間が経過していないか否かの判断を行う。また、7秒間が経過していないと判断された場合は、判断結果がYesとなってステップS g 7に戻る。そして、以下、ステップS g 7、S g 8の処理を繰り返し行い、音声が何等入力されないまま7秒間が経過すると、ステップS g 8の判断結果がNoとなって、ステップS g 9へ進み、音声認識処理を終了してステップS g 1へ戻る。

【0143】一方、音声認識を開始してから、7秒以内に音声が入力されると、ステップS g 7における判断結果がYesとなって、ステップS g 10へ進み、入力された音声の認識を行う。次いで、ステップS g 11へ進んで乱数を発生し、ステップS g 12で所定値 $\alpha$ との比較を行う。ここで、乱数の値が所定値 $\alpha$ 以上であった場合、判断結果がNoとなって図14のステップS g 13へ進み、ステップS g 6で計時が開始されてから7秒間が経過したか否かが判断される。そして、7秒間が経過していなければ、判断結果がYesとなってステップS g 6へ戻り再度音声認識を開始する。また、経過していた場合は判断結果がNoとなってステップS g 1へ戻る。

【0144】一方、図13のステップS g 12において、乱数の値が所定値 $\alpha$ 未満であった場合、判断結果が

Yes となって、ステップ S g 14 において、認識された音声为本実施形態の玩具につけられた名前であった場合、設定モード〔設定 1〕の処理を終了する。また、本実施形態の玩具につけられた名前でなかった場合は、ステップ S g 15 へ進み、ステップ S g 10 で認識された音声に対応するデータ、すなわち、記憶パラメータ MH の値が読み出される。

【0145】そして、ステップ S g 16 に進み、読み出された記憶パラメータ MH の値が 40 以下であった場合、判断結果が No となり、ステップ S g 17 で、ステップ S g 6 における所定時間の計時開始後、7 秒間経過したか否かが判断され、7 秒間経過していなかった場合はステップ S g 6 へ戻り、経過していた場合はステップ S g 1 へ戻る。

【0146】一方、ステップ S g 16 で、読み出された記憶パラメータ MH の値が 40 よりも大きかった場合、判断結果が Yes となり、ステップ S g 18 へ進み、ステップ S g 5 で発声された音声データ（以下、被対応付け音声データという）に対する関連データ RL があるか否かの判断を行う。

【0147】ここで、関連データ RL は、設定モード〔設定 1〕において、被対応付け音声データに、ステップ S g 10 で認識した音声を初めて対応付ける際に、この認識した音声に対して発生されるデータであり、被対応付け音声データが記憶されているアドレスの情報と、被対応付け音声データとの関連度を示す数値と、被対応付け音声データとの対応付けがなされている否かを示すマスタフラグ（初期値 0）からなっている。また、このデータの値が 120 を越えている場合は、被対応付け音声データとの対応付けがなされているものとして、上述したマスタフラグに 1 がセットされる（後述する）。

【0148】そしてステップ S g 18 において、初めて対応付けが行われる場合は、被対応付け音声データに対する関連データ RL が存在していないので、判断結果が Yes となってステップ S g 19 へ進み、被対応付け音声データに対する関連データ RL を発生させる。ここで、関連データ RL の値は、乱数を発生させ、その乱数に、前述した所定値  $\beta$ （図 12、ステップ S f 17 参照）から 10 を引いた値を加えた値とする。

【0149】なお、ステップ S g 19 で発生された関連データ RL は、以後、記憶パラメータ MH と共に記憶され、次回、上記対応付けた音声ステップ S g 10 で認識された場合、ステップ S g 15 において記憶パラメータ MH の値と共にその値が読み出されるものとする。この時、ステップ S g 10 で認識された音声に対して記憶された各関連データ RL のうち、各々に含まれている被対応付け音声データのアドレスを参照して、ステップ S g 5 で発声された音声データのアドレス情報を有する関連データ RL の値が読み出されるものとする。

【0150】また、ステップ S g 18 において、関連デ

ータ RL が既に存在していた場合は、判断結果が No となって、ステップ S g 20 へ進み、関連データ RL の値に 1 が加算される。そして、ステップ S g 19 または S g 20 の処理が終了すると、ステップ S g 21 へ進み、関連データ RL の値が 120 を越えているか否かの判断を行う。関連データ RL の値が 120 以下であったら、判断結果が No となって、ステップ S g 5 で発音された音声データに対し、ステップ S g 10 で認識された音声に対応付けられなかったとして、ステップ S g 1 へ戻る。

【0151】一方、関連データ RL の値が 120 よりも大きかった場合は、両者が対応付けられているとして、判断結果が Yes となり、ステップ S g 22 へ進み、関連データ RL のマスタフラグを 1 にセットする。そして、ステップ S g 23 へ進み、対応付けられた言葉（ステップ S g 5 で発音された言葉）と、対応付けた言葉（ステップ S g 10 で認識された言葉）とを、交互に 5 回ずつ発音させ、ユーザに対して両者が対応付けられたことを通知する。

【0152】このように、上述した設定モードのうち、設定 1 の処理が実行されると、本願実施形態の玩具において音声データ記憶部 15 において学習済みの音声データを、7 秒間隔で順次発声していき、この間にユーザが音声を入力すると、この入力された音声データに対応して記憶され、かつ、上記ユーザが音声を入力する直前に発声された音声データとの関連度を示す関連データ RL が読み出され、そのパラメータの値が 120 を超していれば、両者が対応付けられたものとして、双方の言葉（または音）を交互に 5 回ずつ発声する。

【0153】また、実行モードにおいて、〔設定 1〕で設定した内容を実行する場合、例えば、「ただいま」という言葉が認識されると、「ただいま」という言葉に対して記憶されている各関連データ RL のうち、マスタフラグの値が 1 になっている関連データ RL に含まれているアドレス情報が参照され、そのアドレスに記憶されている音声データが、順次音声データ記憶部 15 から読み出され、スピーカ 20 から発生される。

【0154】なお、上述した関連データ RL の値を、例えば、1 日経過する毎に 1 つデクリメントして、その値が所定値（例えば 0）になった場合、マスタフラグの値を 0 にリセットして対応付けを解消するようにしてもよい。これにより、あたかも、本実施形態における玩具に対し常に言葉の対応を訓練しなければ、覚え込ませた言葉の対応を自然に忘れてしまうかのように振る舞わせることができる。

【0155】さらに、上述した他の実施形態における機能は、先に説明した実施形態と同様、図 10 に示すようなパソコンや、いわゆるテレビゲーム専用機上で実現してもよい。

【0156】

【発明の効果】以上説明したように、請求項1、15、および、17に記載の発明によれば、発せられた音声に対応する音声信号を記憶すると共に、その音声を認識し、認識した音声と記憶した音声信号との対応を設定することができ、また、その設定に基づいて、認識された音声に対応する音声信号を音声に変換して出力するので、単に入力された音声をそのままお返し発生するだけでなく、発せられた音声に対して所望する音声を発生させることができ、よって、発せられた音声に対して返答する音声を多様化させることができる。また、特に人形、または、動物の形状を有するぬいぐるみ、あるいは、室内飾りの内部または外部に設けることにより、飽きにくい玩具を提供することができる。

【0157】また、請求項2に記載の発明によれば、使用者が設定した時刻になると、予め指定した音声を繰り返し発生し、同様に、予め指定した音声を使用者が発声するまで、玩具からの発声が継続されるので、本玩具を目覚まし時計として使用することができる。

【0158】また、請求項3に記載の発明によれば、使用者が設定した音声を発生すると、現在時刻が音声により通知されるので、時刻を知りたい時でも逐一時計の文字盤に視線を移す必要がない。また、使用者が設定した以外の音または音声が発せられた場合は、現在時刻を音声により通知することがないので、例えば雑音等の使用者が意図しない音によって、誤って現在時刻を通知する恐れがなくなる。

【0159】また、請求項4に記載の発明によれば、使用者が設定した時刻に現在時刻を音声により通知するので、通常のタイマとして使用できるばかりでなく、現在の時刻も知ることができる。

【0160】また、請求項5に記載の発明によれば、使用者が設定した時刻になると、同じく使用者によって指定された音声が発生されるので、所定の時刻に使用者が起こすべき行動の内容等を音声で通知することができ、使用者は自らのスケジュールを再確認することができる。

【0161】また、請求項6に記載の発明によれば、入力された音声を任意に記憶し、この任意に記憶した音声を任意に発生するので、画一的な反応を示すことがなくなり、あたかも自律的な意志をもって反応しているかのように振る舞わせることができる。このため、意外性があり、飽きにくい玩具を実現することができる。

【0162】また、請求項7に記載の発明によれば、使用者が発した音声を正確に記憶するか否かが乱数によって左右されるので、正確な音声を記憶するに際し、入力された音声を単純にそのまま記憶するものよりも、人語を習得する動物の、人語を習得する有様をシミュレーションすることができ、意外性が増し、飽きにくい玩具を提供することができる。

【0163】また、請求項8に記載の発明によれば、使

用者が発声した音声を異なる音質で発生するので、人語を習得する動物をシミュレーションすることができ、使用者が発声した音声を自らが聞くといった味気なさを解消することができる。

【0164】また、請求項9に記載の発明によれば、本発明による玩具と同じ玩具から発生された音声が入力された場合は、発生する乱数の数値範囲が変化するので、その様な場合、発生する乱数の数値範囲を狭めるように設定しておけば、玩具同士で音声を記憶させる場合、正確に音声を記憶する割合が増し、同じ玩具同士で互い音声を記憶しあう場合は、容易に正確な音声を記憶できる区別化がなされ、使用者が自らの音声を記憶させる場合とはまた別の楽しみ方が可能となる。

【0165】また、請求項10に記載の発明によれば、同一の音声が発せられた回数以上入力された場合は、発生する乱数の数値範囲が変化するので、例えばその様な場合、発生する乱数の数値範囲を狭めるように設定しておけば、何度同じ音声を入力してもなかなか入力した通りに発生しないといった現象を避けることができる。

【0166】また、請求項11および12に記載の発明によれば、常時計数を行うフリーランカウンタによって計数された値が所定の値と一致した場合、発生する乱数の数値範囲が変化されるので、正確な音声を記憶させるに当たってギャンブル性を加味することができ、使用者の興味を引くことができる。また、発生する乱数の数値範囲が変化された場合は、その旨を所定の音声によって使用者に通知するので、使用者の興味をより一層かき立てることができる。

【0167】また、請求項13に記載の発明によれば、記憶している音声を全て、かつ、順次発生させることができるので、記憶している音声の確認が可能となると共に、相手に宛てたメッセージが記憶されている場合、そのメッセージを確実に相手に伝えることができる。

【0168】また、請求項14に記載の発明によれば、音または音声が発生する際に、人形、または、動物の形状を有するぬいぐるみ玩具本体の一部が動くので、例えば、インコ等のぬいぐるみを本発明の玩具に採用した場合、嘴などが動くようにしておけば、あたかも本物のインコがしゃべっているかのような効果を持たせることができる。

【0169】また、請求項16および18に記載の発明によれば、画像表示装置に人または動物の画像を表示し、音または音声を発生する際に、表示した人または動物の画像が変化するので、視覚的にも楽しむことができる。

【0170】また、請求項19に記載の発明によれば、音声データ出力手段から出力された音声データが一時記憶され、該一時記憶された音声データと同一の音声データが音声データ出力手段から出力された時に乱数が発生され、該乱数と所定値との大小関係が所定の条件を所定

回数満了した時に、1つの音声データが認識された音声に対応付けられる音声データとして記憶され、かつ、上記所定の条件を満たす毎に、画像表示装置に表示させる人または動物の画像に、それぞれ異なる変化が生じるので、ユーザが記憶させようとしている言葉または音が、認識された音声に対応付けられる音声データとして記憶される過程において、どの段階にあるのかを判別することができる。

【0171】また、請求項20に記載の発明によれば、認識された音声に対応付けられる音声データとして音声データを記憶される際に、当該音声データに対し、当該音声データを消去するか否かを決定するための記憶保持値が発生し、該記憶保持値が所定時間が経過する毎に減少し、一定値以下になった時に当該音声データが消去されるので、画像表示装置に表示させる人または動物が、入力頻度の低い言葉または音については、あたかも自然に忘れてしまうかのように振る舞わせることができる。

【0172】また、請求項21に記載の発明によれば、対応付けられる音声データとして記憶された音声データと、認識した音声とを対応づけさせる際に、対応付けられる音声データとして記憶された音声データが、所定時間間隔で順次音声出力手段に出力されると共に、この間に入力された音声を認識し、この認識した音声に対して、その直前に音声出力手段に出力された音声データとの関連度を示す情報が発生し、次回、上記認識した音声、その認識した音声に対して発生された情報によって関連度が示される音声データが音声出力手段に出力された直後に、再度認識された場合、その情報が示す関連度を増加させ、該関連度が所定の条件を満たした時に、上記認識した音声と上記対応付けられる音声データとして記憶させた音声データとを対応付けるので、ある音声に対応付けたい音声を、ユーザ自身が選択して決定キーを押下するという操作を必要とせず、音声の対応付けを行うことができる。

【0173】さらに、請求項22に記載の発明によれば、上記情報が示す関連度が、所定時間経過する毎に減少し、該関連度が一定値以下になった時に、対応付けられた音声と、音声データとの対応を解消するので、あたかも、画像表示装置に表示させる人または動物に対して常に言葉の対応を訓練しなければ、覚え込ませた言葉の対応を自然に忘れてしまうかのように振る舞わせることができる。

#### 【図面の簡単な説明】

【図1】 この発明の一実施形態による音声認識および音声発生装置を備えた玩具の外観を示す正面図である。

【図2】 同音声認識および音声発生装置の構成を示すブロック図である。

【図3】 同音声認識および音声発生装置における学習モードの処理手順を示すフローチャートである。

【図4】 同音声認識および音声発生装置における設定

モードの処理手順を示すフローチャートである。

【図5】 同音声認識および音声発生装置における設定モードの処理手順を示すフローチャートである。

【図6】 同音声認識および音声発生装置における実行モードの処理手順を示すフローチャートである。

【図7】 同音声認識および音声発生装置における実行モードの処理手順を示すフローチャートである。

【図8】 同音声認識および音声発生装置におけるランダムモードの処理手順を示すフローチャートである。

【図9】 同音声認識および音声発生装置におけるメッセージモードの処理手順を示す。

【図10】 同音声認識および音声発生装置にパーソナルコンピュータを用いた場合の構成を示すブロック図である。

【図11】 同音声認識および音声発生装置における学習モードの、他の実施形態の処理手順を示すフローチャートである。

【図12】 同音声認識および音声発生装置における学習モードの、他の実施形態の処理手順を示すフローチャートである。

【図13】 同音声認識および音声発生装置における設定モード・設定1の、他の実施形態の処理手順を示すフローチャートである。

【図14】 同音声認識および音声発生装置における設定モード・設定1の、他の実施形態の処理手順を示すフローチャートである。

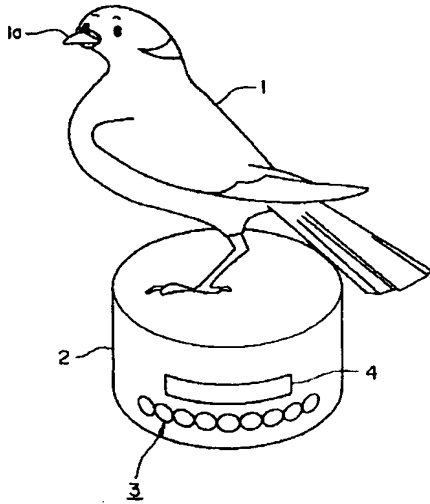
#### 【符号の説明】

1	ぬいぐるみ
1 a	嘴
2	本体
3	操作部
4	表示部
10	マイクロフォン
11, 18	増幅回路
12	フィルタ
13	A/D変換器
14	音声データ変更部
15	音声データ記憶部
16	D/A変換器
17	音声変換回路
19	音量調整回路
20	スピーカ
21	特徴パターン抽出部
22	音声入力検出部
23	タイマ
24	玩具特徴パターン記憶部
25	特徴パターン記憶部
26	カウンタ
27	乱数発生部
28	駆動装置

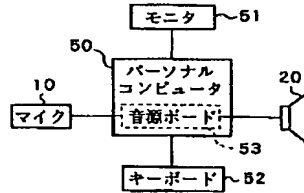
29 制御部  
50 パーソナルコンピュータ  
51 モニタ

\* 52 キーボード  
53 音源ボード  
\*

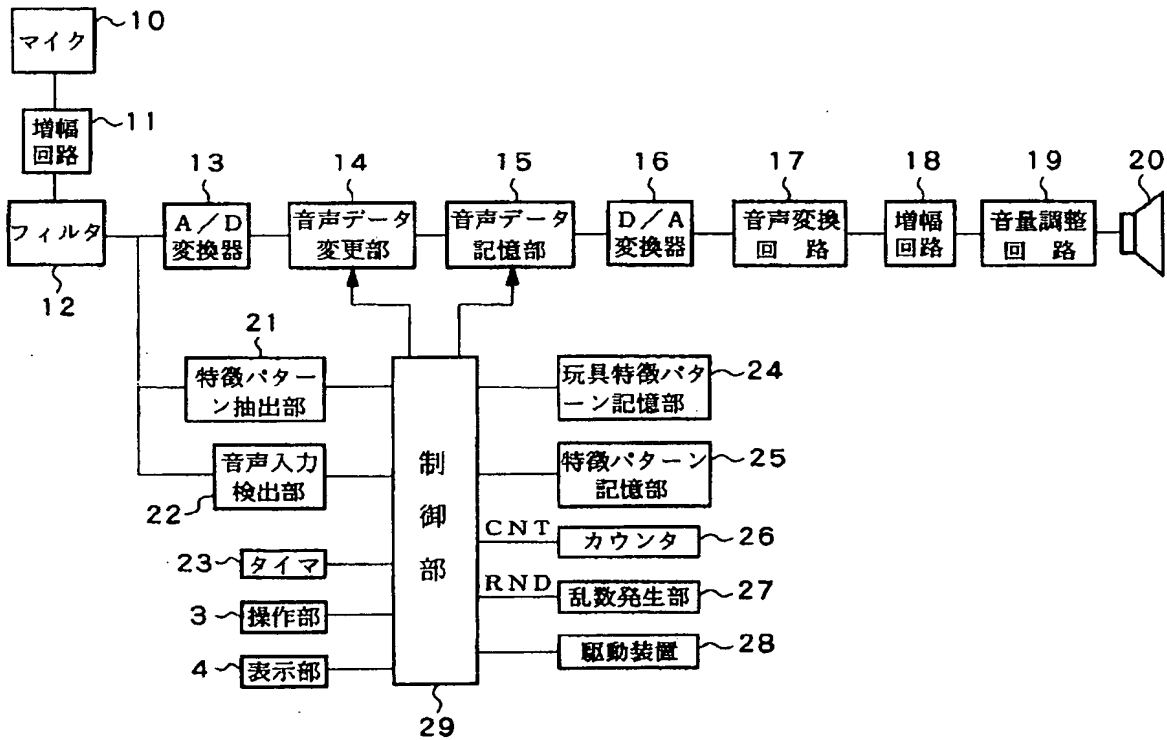
【図1】



【図10】

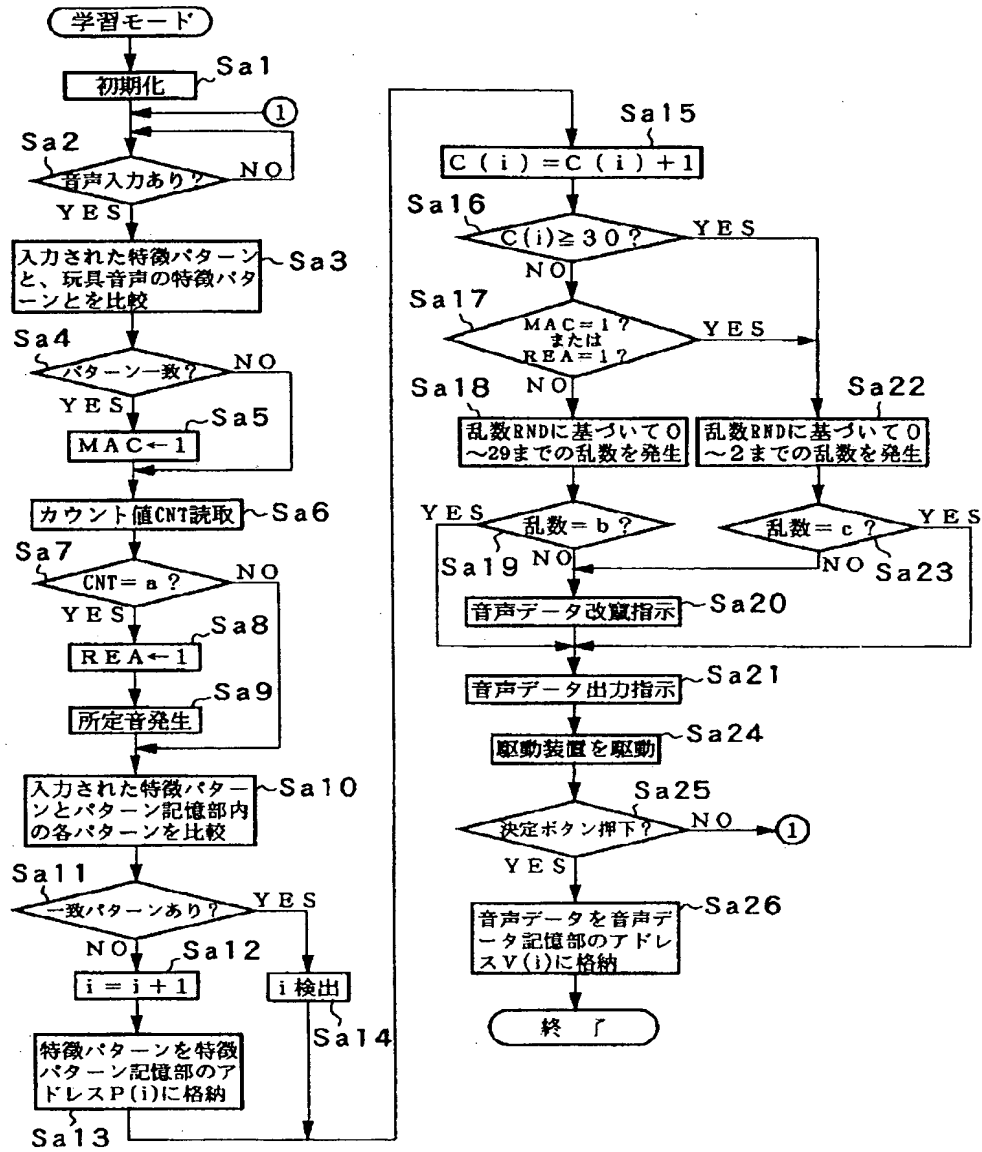


【図2】

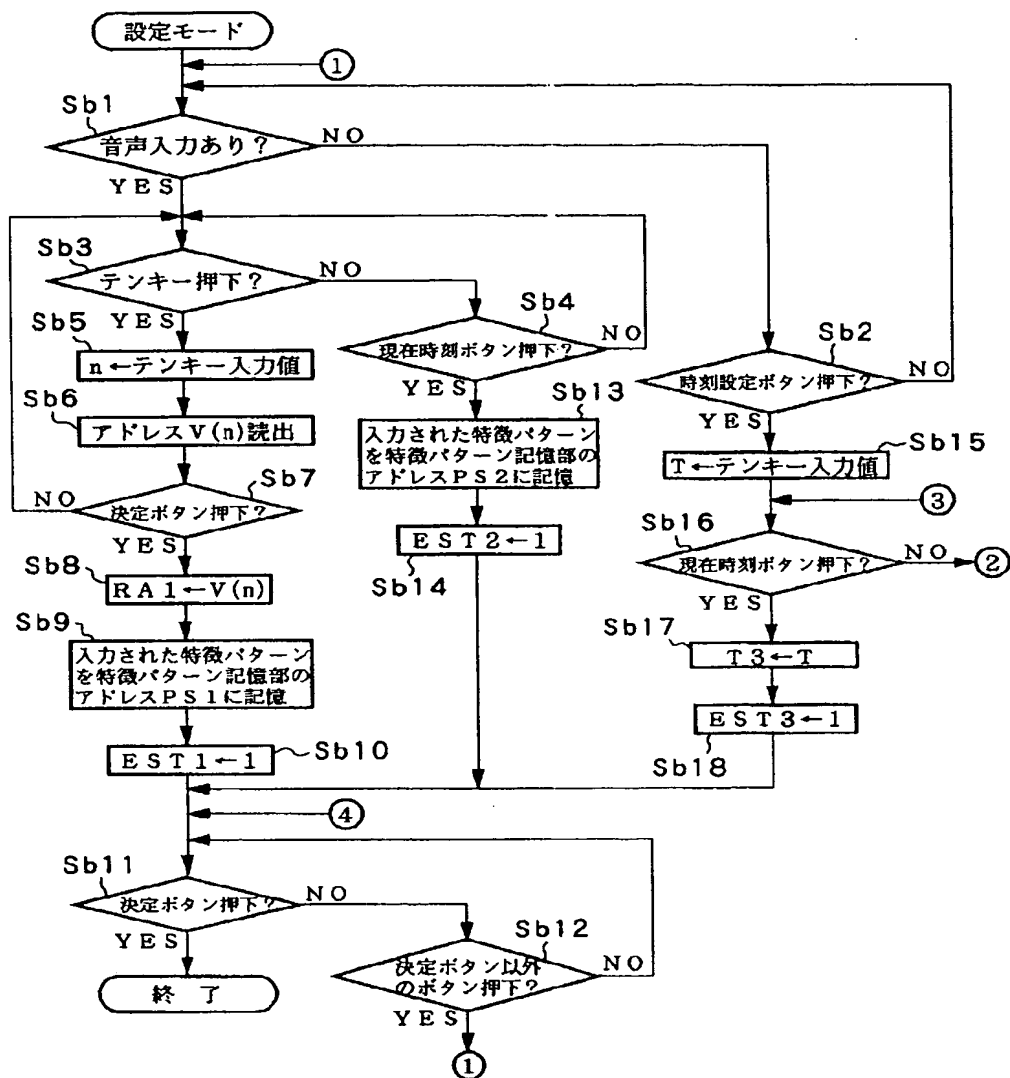




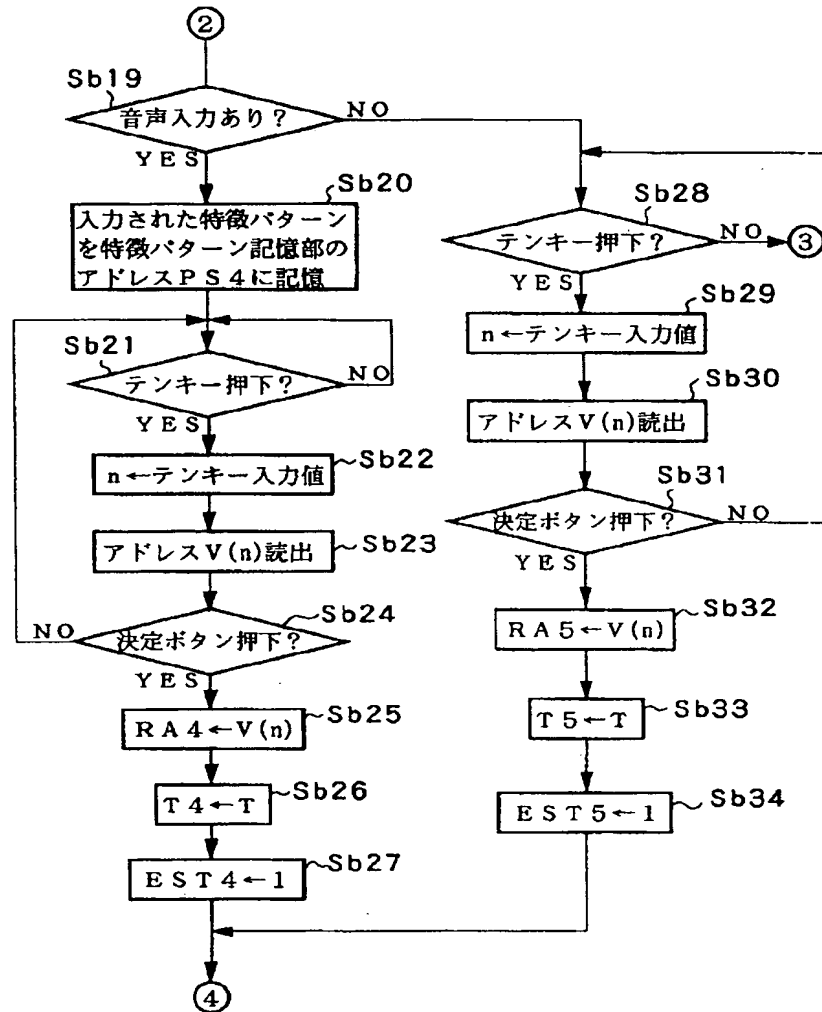
【図3】



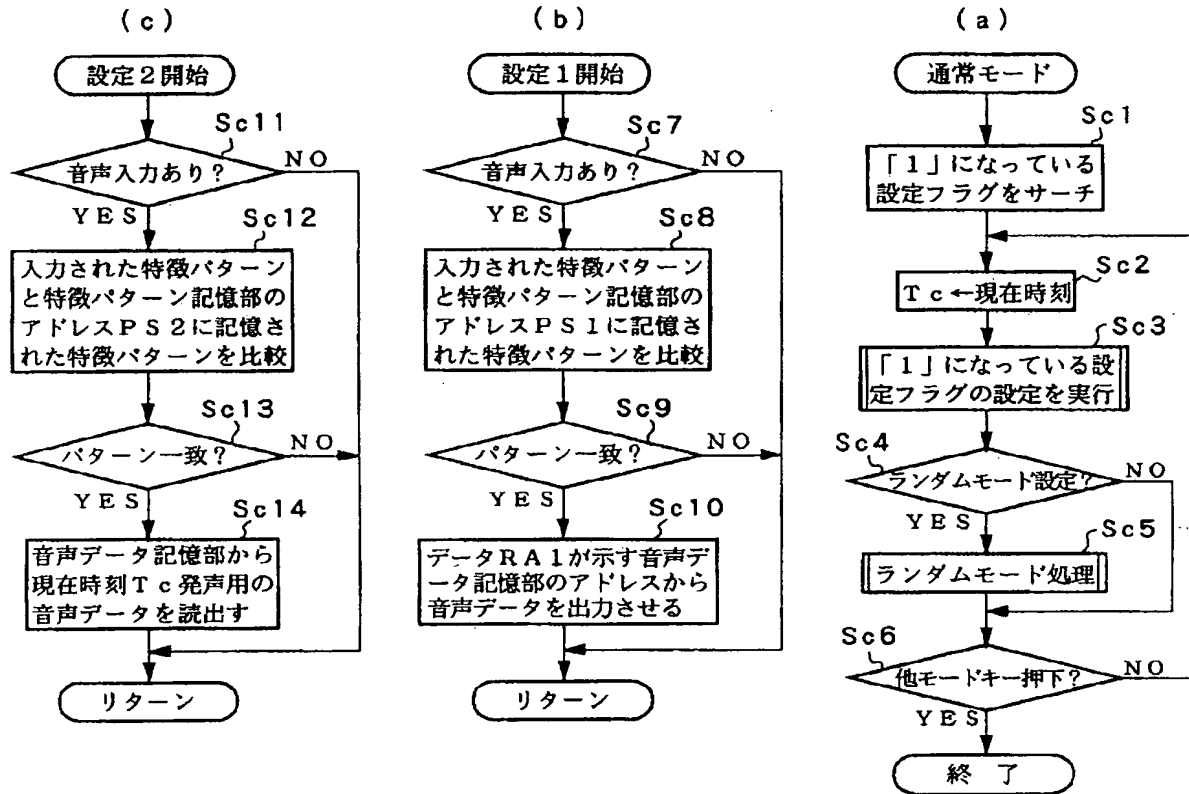
【図4】



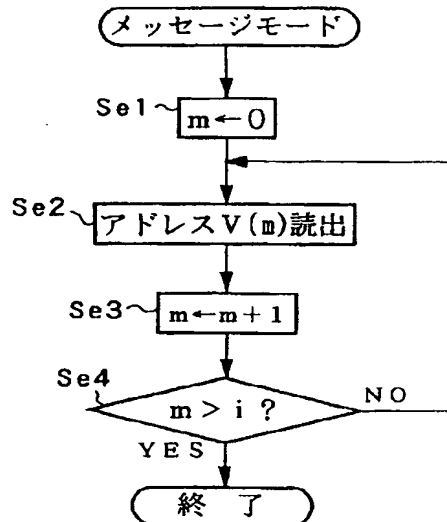
【図5】



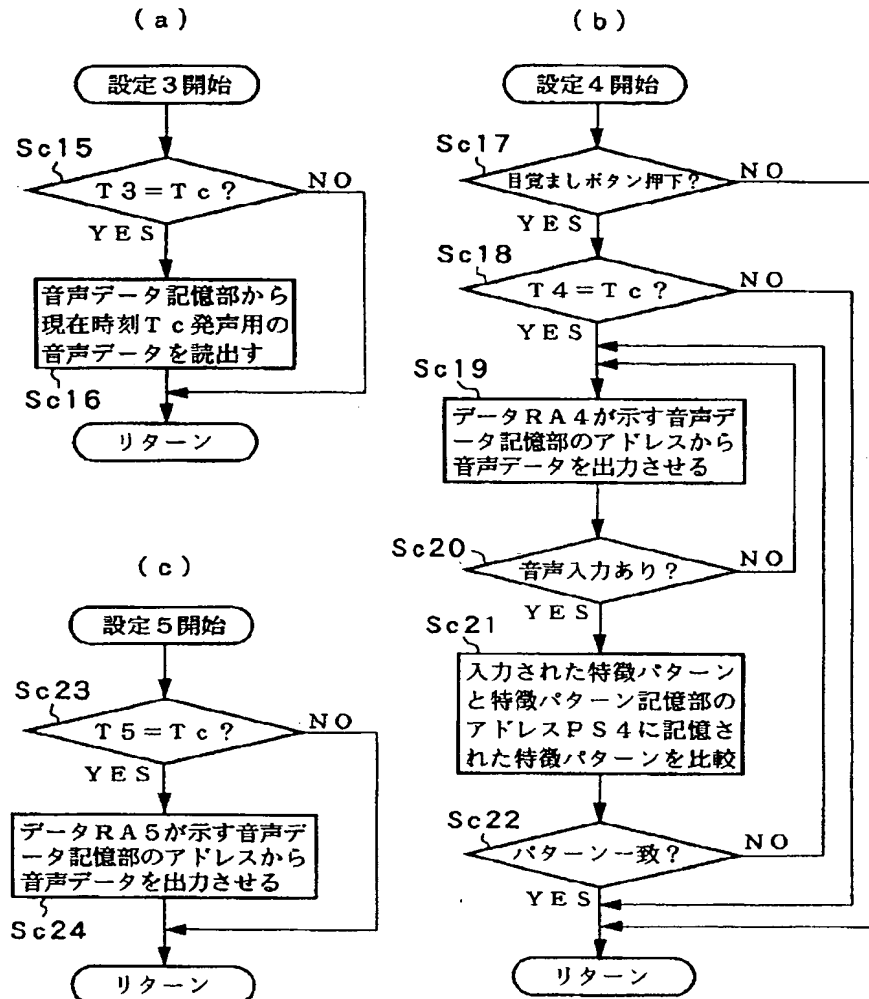
【図6】



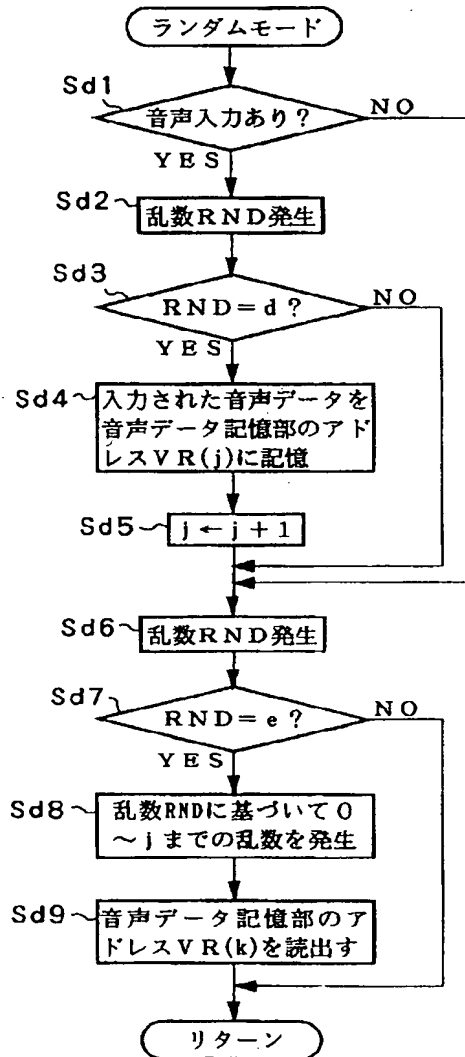
【図9】



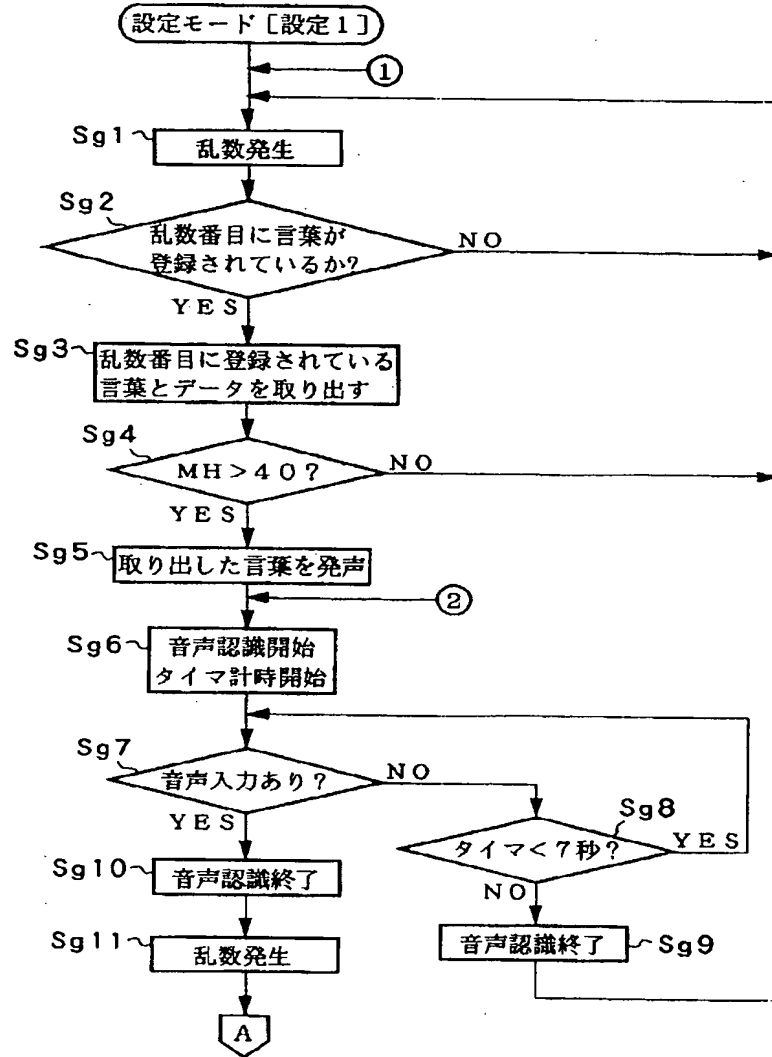
【図7】



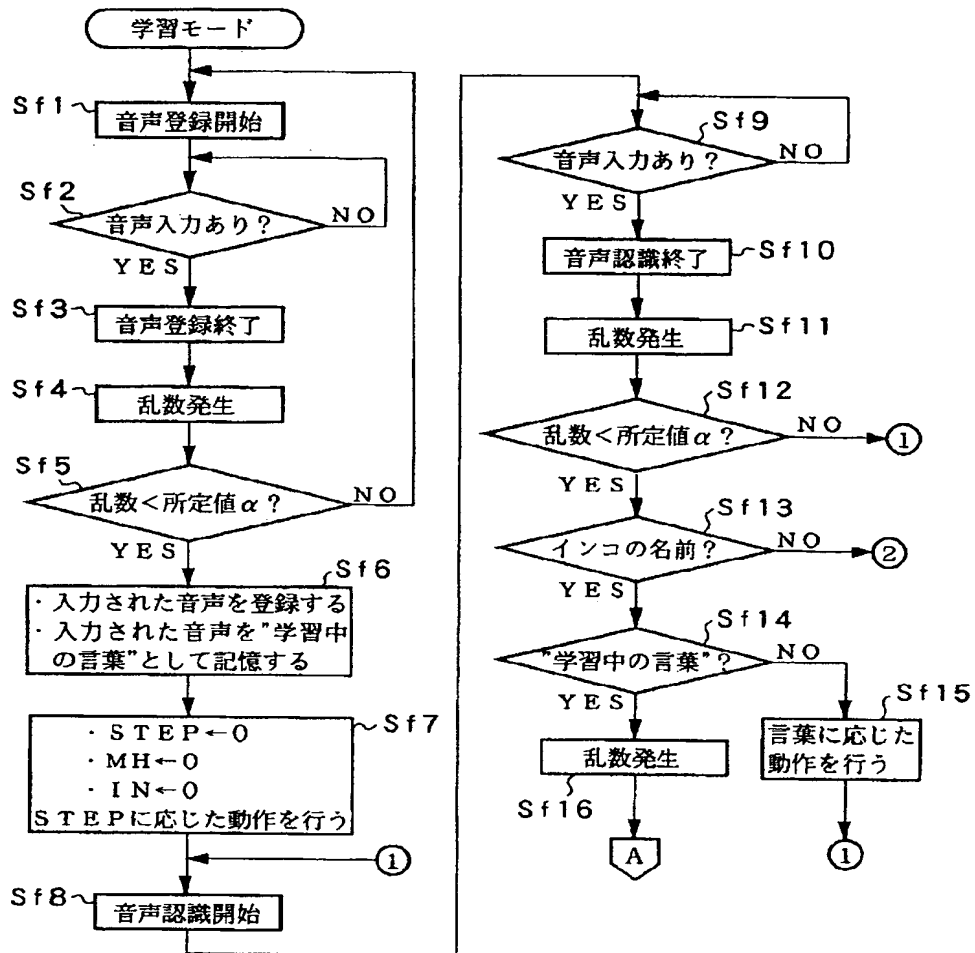
【図8】



【図13】

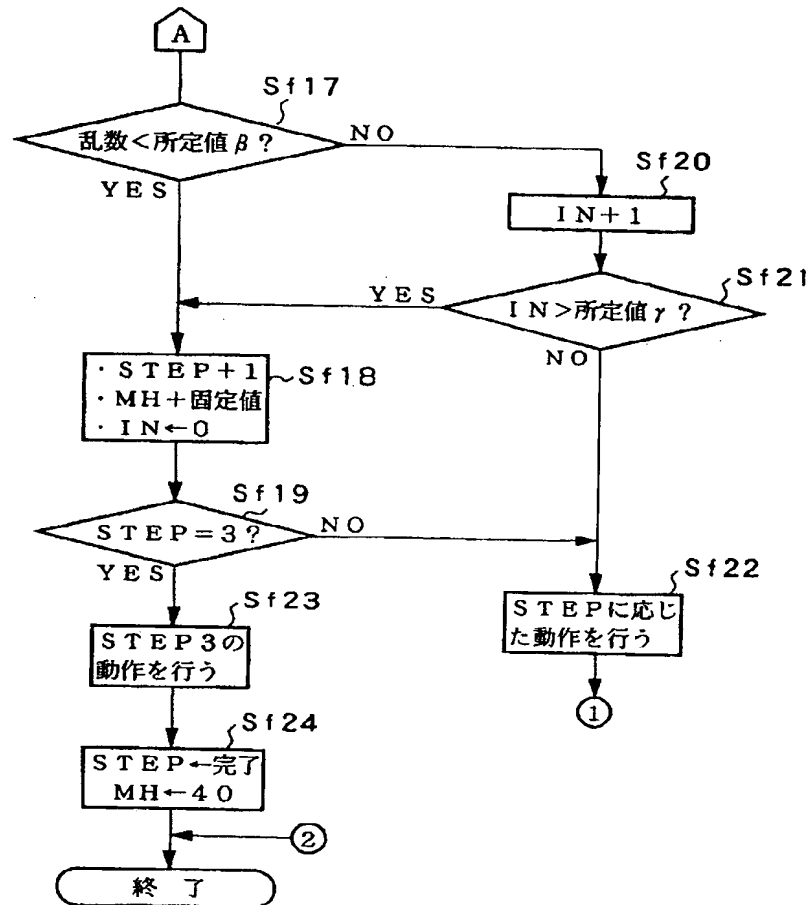


【図11】

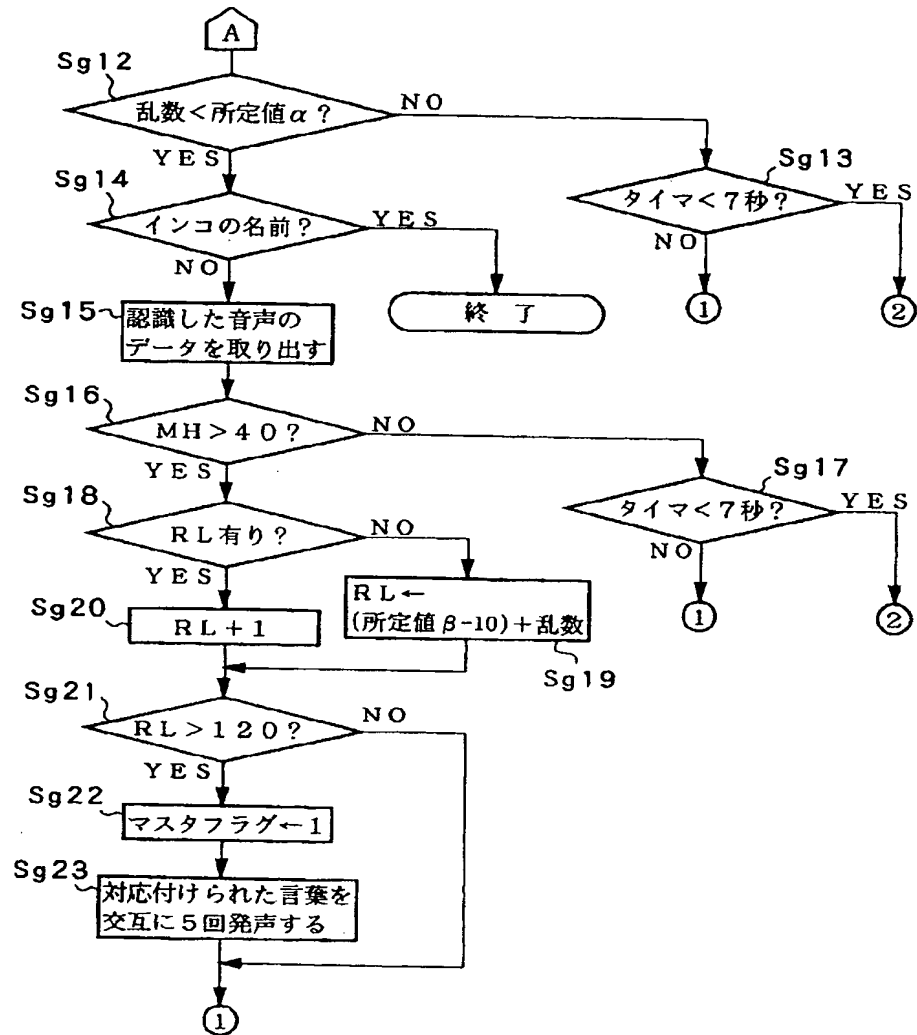




【図12】



【図14】



# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-179941

(43)Date of publication of application : 07.07.1998

---

(51)Int.Cl. A63H 3/33

A63H 5/00

G10L 3/00

G10L 3/00

---

(21)Application number : 09-262557 (71)Applicant : KUWABARA JUNJI

TAKAYANAGI TAKASHI

(22)Date of filing : 26.09.1997 (72)Inventor : KUWABARA JUNJI

---

(30)Priority

Priority number : 08278353

Priority date : 21.10.1996

Priority country : JP

---

(54) VOICE RECOGNITION AND GENERATION APPARATUSES, TOY WITH  
THE APPARATUS, AND RECORDING MEDIUM HAVING VOICE  
RECOGNITION AND GENERATION CONTROL PROGRAM RECORDED  
THEREIN

(57)Abstract:

PROBLEM TO BE SOLVED: To provide voice recognition and voice generation  
apparatuses which enable optionally setting of correspondence of a voice  
inputted to a voice to be generated for this voice, a toy with the same and a  
recording medium having voice recognition and generation control programs  
recorded therein.

SOLUTION: A sound inputted from a microphone 10 is converted to a voice data by an A/D converter 13 and finally stored into a voice data memory section 15. A feature pattern of the voice inputted from a separated microphone 10 is extracted by a feature pattern extracting part 21, the extracted feature pattern is made to correspond to any of voice data stored in the voice data memory section 15 based on a direction from an operating section 3 and then, stored into the feature pattern memory section 25. As the result, when the feature pattern stored in the feature pattern memory section 25 is extracted again, a voice data corresponding to the feature pattern is read out of the voice data memory section 15 and finally emitted as sound from a speaker 20.

---

LEGAL STATUS [Date of request for examination] 26.09.1997

[Date of sending the examiner's decision of rejection] 22.08.2000

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

**\* NOTICES \***

JP0 and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.

2. \*\*\*\* shows the word which can not be translated.

3. In the drawings, any words are not translated.

---

**CLAIMS**

---

[Claim(s)]

[Claim 1] A sound signal output means to output the sound signal corresponding to the inputted voice, A sound signal storage means to memorize the sound signal outputted from said sound signal output means, A speech recognition means to recognize the voice inputted into said sound signal output means based on the sound signal outputted from said sound signal output means, With

a correspondence setting means to set up correspondence with the sound signal memorized by said sound signal storage means and the voice recognized by said speech recognition means, and said speech recognition means The sound signal read-out means which will read the sound signal corresponding to the recognized voice from said sound signal storage means according to the correspondence set up by said correspondence setting means if said matched voice is recognized, The speech recognition and the speech generation device possessing a voice output means to change and output the sound signal which said sound signal read-out means read to voice The toy equipped with the speech recognition and the speech generation device which are characterized by preparing in the interior or the exterior of sewing-basis which has the configuration of a doll or an animal, and indoor decoration.

[Claim 2] a time designated means to specify time of day, and the time check which clocks current time -- a means and said time check, if a means clocks the time of day specified by said time designated means The repetitive read-out means which repeats and reads the sound signal memorized by said sound signal storage means, If the specific voice corresponding to the sound signal repeatedly read by said repetitive read-out means is recognized by said speech recognition means according to the correspondence set up by said correspondence setting means The toy equipped with the speech recognition



and the speech generation device according to claim 1 which are characterized by changing and outputting to voice the sound signal from which it comes to have the read-out means for stopping which stops read-out of the sound signal by said repetitive read-out means, and said voice output means was read by said repetitive read-out means.

[Claim 3] said time check -- the toy equipped with the speech recognition and the speech generation device according to claim 1 or 2 which are characterized by providing a notice means of current time to notify with voice the current time clocked by the means, and the notice control means of time of day controlled so that said notice means of current time notifies current time when specific voice has been recognized by said speech recognition means.

[Claim 4] said notice control means of time of day -- said time check -- the toy equipped with the speech recognition and the speech generation device according to claim 3 which are characterized by controlling so that said notice means of current time notifies current time when a means clocks the time of day specified by said time designated means.

[Claim 5] A selection means to choose any one of the sound signals memorized by said sound signal storage means, said time check, when a means clocks the time of day specified by said time designated means It comes to have the appointed time-of-day read-out means which reads the sound signal chosen by

said selection means from said sound signal storage means. The toy equipped with the speech recognition and the speech generation device according to claim 1 which are characterized by said voice output means changing and outputting to voice the sound signal read by said appointed time-of-day read-out means.

[Claim 6] A storage control means to make arbitration memorize the sound signal outputted from said sound signal output means for said sound signal storage means, It has the arbitration read-out means which reads the sound signal of arbitration to said sound signal storage means among the sound signals memorized by arbitration. Said voice output means The toy which equipped any 1 term with the speech recognition and the speech generation device of a publication claim 1 characterized by changing and outputting to voice the sound signal which was read by said arbitration read-out means, and which was memorized by arbitration thru/or among 5.

[Claim 7] The random number generated from a random-number-generation means to generate a random number, and said random-number-generation means When not in agreement with the numeric value set up beforehand, it comes to have a sound signal modification means to change and output the information included in the sound signal outputted from said sound signal output means. The toy which equipped any 1 term with the speech recognition and the

speech generation device of a publication claim 1 characterized by said voice output means changing and outputting the sound signal outputted from said sound signal modification means to voice thru/or among 6.

[Claim 8] The toy equipped with the speech recognition and the speech generation device according to claim 7 which are characterized by providing a tone-quality modification means to change the tone quality of the voice outputted from said voice output means.

[Claim 9] The toy equipped with the speech recognition and the speech generation device according to claim 8 which are characterized by providing the random-number-generation control means to which the numerical range of the random number which said random-number-generation means generates is changed when the sound signal outputted from said sound signal output means is a sound signal which had tone quality changed by said tone-quality modification means.

[Claim 10] The toy equipped with the speech recognition and the speech generation device according to claim 8 which are characterized by providing the random-number-generation control means to which the numerical range of the random number which said random-number-generation means generates is changed when the same voice has been recognized more than the count set up beforehand by said speech recognition means.

[Claim 11] It is the toy equipped with the speech recognition and the speech generation device according to claim 8 which are characterized by providing the random-number-generation control means to which the numerical range of the random number which said random-number-generation means generates is changed when a sound signal is outputted from the counter which always carries out counting in the predetermined numerical range, and said sound signal output means and the enumerated data are [ the enumerated data of said counter are read and ] in agreement with a predetermined numeric value.

[Claim 12] The toy equipped with the speech recognition and the speech generation device according to claim 11 which are characterized by to have the predetermined sound signal read-out means which reads said predetermined sound signal from said predetermined sound signal storage means, and for said voice output means to change and output to voice the predetermined sound signal read by said predetermined sound signal read-out means when a predetermined sound signal storage means to by\_which the predetermined sound signal was memorized, the enumerated data of said counter, and said predetermined numeric value are in agreement.

[Claim 13] The toy which equipped any 1 term with the speech recognition and the speech generation device of a publication claim 1 characterized by changing and outputting to voice the sound signal from which it comes to have the

sequential read-out means which reads altogether the sound signal memorized by said sound signal storage means one by one, and said voice output means was read by said sequential read-out means thru/or among 11.

[Claim 14] Sewing-basis and the indoor decoration which have the configuration of said doll or an animal are the toy which equipped any 1 term with the speech recognition and the speech generation device of a publication claim 1 characterized by providing the driving means which makes the moving part drive when it has moving part and voice is outputted from said voice output means thru/or among 12.

[Claim 15] A voice data output means to output the voice data corresponding to the inputted voice, A voice data storage means to memorize the voice data outputted from said voice data output means, A speech recognition means to recognize the voice inputted into said voice data output means based on the voice data outputted from said voice data output means, A correspondence setting means to set up correspondence with the voice data memorized by said voice data storage means and the voice recognized by said speech recognition means, If the voice matched by said correspondence setting means is recognized by the speech recognition means The voice data read-out means which reads the voice data corresponding to said recognized voice from said voice data storage means according to the correspondence set up by said

correspondence setting means, The speech recognition and the speech generation device which are characterized by providing a voice output means to change and output the voice data which said voice data read-out means read to voice and which perform speech recognition and voice generating by computer.

[Claim 16] The speech recognition and the speech generation device which perform speech recognition and voice generating by computer according to claim 15 characterized by having the image control means controlled to change the image of the man or animal displayed on said image display means when an image display means to display the image of a man or an animal, and said voice data read-out means read voice data from said voice data storage means.

[Claim 17] A voice data output means to output the voice data corresponding to the inputted voice, It is the record medium which recorded the program for making speech recognition and voice generating perform on the computer which has a voice output means to change and output the voice data corresponding to voice to voice. This program makes the voice data outputted from said voice data output means memorize, and said inputted voice is made to recognize based on the voice data outputted from said voice data output means. Said voice data made to memorize, The record medium which recorded the speech recognition and the voice generating control program which are characterized by making the voice data corresponding to the voice output to said voice output

means out of said voice data made to memorize when said recognized voice is made to be matched and said voice made to be matched has been recognized again.

[Claim 18] Said program is the record medium which recorded the speech recognition according to claim 17 and the voice generating control program which are characterized by changing the image of the person who made it display on said image display device, or an animal to it in case said computer is made to display the image of a man or an animal on an image display device and is made to output the voice data corresponding to the voice to said voice output means out of said voice data made to memorize.

[Claim 19] Said program the voice data outputted to said computer from said voice data output means In case it is made to memorize as voice data matched with said recognized voice Store temporarily the voice data outputted from said voice data output means, and a random number is generated when the same voice data as the voice data made to this store temporarily is outputted from said voice data output means. Whenever the size relation between this random number and a predetermined value fulfills predetermined conditions, in the image of those who make it display on said image display device, or an animal The record medium which recorded the speech recognition according to claim 18 and the voice generating control program which are characterized by making

the voice data concerned memorize as voice data matched with said recognized voice when a change different, respectively is produced and said predetermined conditions are fulfilled the number of predetermined times.

[Claim 20] Said program is the record medium which recorded the speech recognition according to claim 18 and the voice generating control program which are characterized by to eliminate the voice data concerned when the storage retention data for determining whether eliminate the voice data concerned to the voice data concerned in case voice data is made to memorize as voice data matched with said recognized voice were generated, these storage retention data are decreased whenever predetermined time passed, and it becomes below constant value.

[Claim 21] The voice data which stored said program in said computer as said voice data matched, In case said recognized voice is made to be matched, the voice data made to memorize as said voice data matched is made to output to the sequential aforementioned voice output means at intervals of predetermined time. As opposed to the voice which has recognized and this recognized the voice inputted while making the voice data made to memorize as said voice data matched output to a voice output means at intervals of said predetermined time The information which shows degree of association with the voice data which said voice output means was made to output just before that is generated. The



recognized this voice immediately after outputting the voice data in which relation is shown using the information generated to the this recognized voice to said voice output means When it has recognized again, and the degree of association which the information shows is made to increase and this degree of association fulfills predetermined conditions The record medium which recorded the speech recognition according to claim 19 and the voice generating control program which are characterized by matching the voice data made to memorize as said recognized voice and said voice data matched.

[Claim 22] Said program is the record medium which recorded the speech recognition according to claim 21 and the voice generating control program which are characterized by canceling correspondence with the matched voice and voice data when the degree of association which said information shows was decreased and this degree of association became below constant value, whenever predetermined time passed.

---

## DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates to the record medium which recorded the toy equipped with the speech recognition which utters the voice corresponding to the inputted voice, a speech generation device, and this equipment, speech recognition, and a voice generating control program.

[0002]

[Description of the Prior Art] While memorizing the inputted voice conventionally and pronouncing the memorized voice as it is, the toy which reproduces the sound recorded if needed is devised by having a playback switch etc. and operating the playback switch. For example, if sewing-basis of a parrot is equipped with a microphone, an A-D converter, semiconductor memory, a DA converter, a loudspeaker, etc., and voice generated toward the microphone is digital-signal-ized by an A-D converter etc., it records on semiconductor memory as voice data and the sound signal from a microphone is lost, voice data is read from semiconductor memory one by one, and it outputs to a DA converter, and is made to generate as voice from a loudspeaker in JP,63-290594,A. Moreover, in the toy mentioned above, a playback switch etc. is formed and it has the function to reproduce it when asking for the voice data memorized by semiconductor memory.

[0003]

[Problem(s) to be Solved by the Invention] By the way, in the toy mentioned

above, since the inputted voice was memorized correctly and was always pronounced, it does not have unexpected nature in any way, and it may have got bored with it immediately. Moreover, since the inputted voice was only uttered as it was, it was restricted to the application as a recorded message sender for telephone which tells a mere toy and a mere partner a message.

[0004] This invention was made in view of the situation mentioned above, and aims at offering the record medium which recorded the toy equipped with the speech recognition which can set correspondence with the voice inputted from the outside, and the voice uttered to that voice as arbitration, a speech generation device, and its equipment, speech recognition, and a voice generating control program. Moreover, as other purposes, the application does not remain only in a toy or a recorded message sender for telephone, but it has the notice function of a schedule and the alarm clock function which notify with voice the contents of the action which a user should take at predetermined time of day etc., and aims at offering the record medium which recorded the toy equipped with the speech recognition and the speech generation device which can respond to various applications, and its equipment, speech recognition, and a voice generating control program.

[0005]

[Means for Solving the Problem] A sound signal output means to output the

sound signal corresponding to the voice into which invention according to claim 1 was inputted, A sound signal storage means to memorize the sound signal outputted from said sound signal output means, A speech recognition means to recognize the voice inputted into said sound signal output means based on the sound signal outputted from said sound signal output means, With a correspondence setting means to set up correspondence with the sound signal memorized by said sound signal storage means and the voice recognized by said speech recognition means, and said speech recognition means The sound signal read-out means which will read the sound signal corresponding to the recognized voice from said sound signal storage means according to the correspondence set up by said correspondence setting means if said matched voice is recognized, The speech recognition and the speech generation device possessing a voice output means to change and output the sound signal which said sound signal read-out means read to voice It is the toy equipped with the speech recognition and the speech generation device which are characterized by preparing in the interior or the exterior of sewing-basis which has the configuration of a doll or an animal, and indoor decoration.

[0006] In the toy with which invention according to claim 2 was equipped with speech recognition and a speech generation device according to claim 1 a time designated means to specify time of day, and the time check which clocks

current time -- a means and said time check, if a means clocks the time of day specified by said time designated means The repetitive read-out means which repeats and reads the sound signal memorized by said sound signal storage means, If the specific voice corresponding to the sound signal repeatedly read by said repetitive read-out means is recognized by said speech recognition means according to the correspondence set up by said correspondence setting means It is characterized by changing and outputting to voice the sound signal from which it comes to have the read-out means for stopping which stops read-out of the sound signal by said repetitive read-out means, and said voice output means was read by said repetitive read-out means.

[0007] the toy with which invention according to claim 3 was equipped with speech recognition and a speech generation device according to claim 1 or 2 -- setting -- said time check -- it is characterized by providing a notice means of current time to notify with voice the current time clocked by the means, and the notice control means of time of day controlled so that said notice means of current time notifies current time when specific voice has been recognized by said speech recognition means.

[0008] the toy with which invention according to claim 4 was equipped with speech recognition and a speech generation device according to claim 3 -- setting -- said notice control means of time of day -- said time check -- when a

means clocks the time of day specified by said time designated means, it is characterized by controlling so that said notice means of current time notifies current time.

[0009] In the toy with which invention according to claim 5 equipped any 1 term with the speech recognition and the speech generation device of a publication claim 2 thru/or among 4 A selection means to choose any one of the sound signals memorized by said sound signal storage means, said time check, when a means clocks the time of day specified by said time designated means It is characterized by changing and outputting to voice the sound signal from which it comes to have the appointed time-of-day read-out means which reads the sound signal chosen by said selection means from said sound signal storage means, and said voice output means was read by said appointed time-of-day read-out means.

[0010] In the toy with which invention according to claim 6 equipped any 1 term with the speech recognition and the speech generation device of a publication claim 1 thru/or among 5 A storage control means to make arbitration memorize the sound signal outputted from said sound signal output means for said sound signal storage means, It has the arbitration read-out means which reads the sound signal of arbitration to said sound signal storage means among the sound signals memorized by arbitration, and is characterized by said voice output

means changing and outputting to voice the sound signal which was read by said arbitration read-out means and which was memorized by arbitration.

[0011] In the toy with which invention according to claim 7 equipped any 1 term with the speech recognition and the speech generation device of a publication claim 1 thru/or among 6 The random number generated from a random-number-generation means to generate a random number, and said random-number-generation means When not in agreement with the numeric value set up beforehand, it comes to have a sound signal modification means to change and output the information included in the sound signal outputted from said sound signal output means. It is characterized by said voice output means changing and outputting the sound signal outputted from said sound signal modification means to voice.

[0012] Invention according to claim 8 is characterized by providing a tone-quality modification means to change the tone quality of the voice outputted from said voice output means in the toy equipped with speech recognition and a speech generation device according to claim 7.

[0013] In the toy equipped with speech recognition and a speech generation device according to claim 8, invention according to claim 9 is characterized by providing the random-number-generation control means to which the numerical range of the random number which said random-number-generation means

generates is changed, when the sound signal outputted from said sound signal output means is a sound signal which had tone quality changed by said tone-quality modification means.

[0014] In the toy equipped with speech recognition and a speech generation device according to claim 8, invention according to claim 10 is characterized by providing the random-number-generation control means to which the numerical range of the random number which said random-number-generation means generates is changed, when the sound signal outputted from said sound signal output means is a sound signal which had tone quality changed by said tone-quality modification means.

[0015] In the toy with which invention according to claim 11 was equipped with speech recognition and a speech generation device according to claim 8 When a sound signal is outputted from the free run counter which always carries out counting in the predetermined numerical range, and said sound signal output means, When the enumerated data of said free run counter are read and the enumerated data are in agreement with a predetermined numeric value, it is characterized by providing the random-number-generation control means to which the numerical range of the random number which said random-number-generation means generates is changed.

[0016] In the toy with which invention according to claim 12 was equipped with



speech recognition and a speech generation device according to claim 11 A predetermined sound signal storage means by which the predetermined sound signal was memorized, and the enumerated data of said free run counter, When said predetermined numeric value is in agreement, it has the predetermined sound signal read-out means which reads said predetermined sound signal from said predetermined sound signal storage means. It is characterized by said voice output means changing and outputting to voice the predetermined sound signal read by said predetermined sound signal read-out means.

[0017] Invention according to claim 13 is characterized by changing and outputting to voice the sound signal from which it comes to have the sequential read-out means which reads altogether the sound signal memorized by said sound signal storage means one by one, and said voice output means was read by said sequential read-out means in the toy which equipped any 1 term with the speech recognition and the speech generation device of a publication claim 1 thru/or among 12.

[0018] In the toy with which invention according to claim 14 equipped any 1 term with the speech recognition and the speech generation device of a publication claim 1 thru/or among 13, it is characterized by sewing-basis and the indoor decoration which have the configuration of said doll or an animal possessing the driving means which makes the moving part drive, when it has moving part and

voice is outputted from said voice output means.

[0019] A voice data output means to output the voice data corresponding to the voice into which invention according to claim 15 was inputted, A voice data storage means to memorize the voice data outputted from said voice data output means, A speech recognition means to recognize the voice inputted into said voice data output means based on the voice data outputted from said voice data output means, A correspondence setting means to set up correspondence with the voice data memorized by said voice data storage means and the voice recognized by said speech recognition means, If the voice matched by said correspondence setting means is recognized by the speech recognition means The voice data read-out means which reads the voice data corresponding to said recognized voice from said voice data storage means according to the correspondence set up by said correspondence setting means, It is the speech recognition and the speech generation device which are characterized by providing a voice output means to change and output the voice data which said voice data read-out means read to voice and which perform speech recognition and voice generating by computer.

[0020] In speech recognition and a speech generation device according to claim 15, invention according to claim 16 is characterized by having the image control means which controls the image of the man or animal displayed on said image

display means to change, when an image display means to display the image of a man or an animal, and said voice data read-out means read voice data from said voice data storage means.

[0021] A voice data output means to output the voice data corresponding to the voice into which invention according to claim 17 was inputted, It is the record medium which recorded the program for making speech recognition and voice generating perform on the computer which has a voice output means to change and output the voice data corresponding to voice to voice. This program makes the voice data outputted to said computer from said voice data output means memorize. Said inputted voice is made to recognize based on the voice data outputted from said voice data output means. Said voice data made to memorize, When said recognized voice is made to be matched and said voice made to be matched has been recognized again, it is characterized by making the voice data corresponding to the voice output to said voice output means out of said voice data made to memorize.

[0022] In the record medium with which invention according to claim 18 recorded speech recognition according to claim 17 and a voice generating control program An image display device is made to display the image of a man or an animal on said computer. In case the voice data corresponding to the voice is made to output to said voice output means out of said voice data made to

memorize, it is characterized by recording the program to which the image of the person who made it display on said image display device, or an animal is changed.

[0023] In the record medium with which invention according to claim 19 recorded speech recognition according to claim 18 and a voice generating control program Said program the voice data outputted to said computer from said voice data output means In case it is made to memorize as voice data matched with said recognized voice Store temporarily the voice data outputted from said voice data output means, and a random number is generated when the same voice data as the voice data made to this store temporarily is outputted from said voice data output means. Whenever the size relation between this random number and a predetermined value fulfills predetermined conditions, in the image of those who make it display on said image display device, or an animal When a change different, respectively is produced and said predetermined conditions are fulfilled the number of predetermined times, it is characterized by making the voice data concerned memorize as voice data matched with said recognized voice.

[0024] In the record medium with which invention according to claim 20 recorded speech recognition according to claim 18 and a voice generating control program In case said program makes voice data memorize as voice data

matched with said recognized voice When the storage retention data for determining whether eliminate the voice data concerned to the voice data concerned were generated, these storage retention data are decreased whenever predetermined time passed, and it becomes below constant value, it is characterized by eliminating the voice data concerned.

[0025] In the record medium with which invention according to claim 21 recorded speech recognition according to claim 19 and a voice generating control program The voice data which said program stored in said computer as said voice data matched, In case said recognized voice is made to be matched, the voice data made to memorize as said voice data matched is made to output to the sequential aforementioned voice output means at intervals of predetermined time. As opposed to the voice which has recognized and this recognized the voice inputted while making the voice data made to memorize as said voice data matched output to a voice output means at intervals of said predetermined time The information which shows degree of association with the voice data which said voice output means was made to output just before that is generated. The recognized this voice immediately after outputting the voice data in which relation is shown using the information generated to the this recognized voice to said voice output means When it has recognized again, and the degree of association which the information shows is made to increase and this degree of

association fulfills predetermined conditions, it is characterized by matching the voice data made to memorize as said recognized voice and said voice data matched.

[0026] In the record medium which recorded speech recognition according to claim 21 and a voice generating control program, invention according to claim 22 is characterized by canceling correspondence with the matched voice and voice data, when said program decreases the degree of association which said information shows whenever predetermined time passed, and this degree of association becomes below constant value.

[0027]

[Embodiment of the Invention] Hereafter, 1 operation gestalt of this invention is explained with reference to a drawing. the microphone (illustration abbreviation) which changes into an electrical signal the voice which drawing 1 is drawing showing the appearance of the toy (only henceforth a toy) equipped with the speech recognition and the speech generation device in this operation gestalt, and 1 is sewing-basis which imitated the parakeet in this drawing, and was inputted into that head -- moreover, the loudspeaker (illustration abbreviation) which utters voice is built in the throat, respectively. Moreover, 1a is a sewing [ which imitated the parakeet ]-based beak, and when voice is uttered from a loudspeaker, it opens it and closes as if the sewing-based parakeet had talked.

[0028] 2 is the body section of the toy in this operation gestalt, and if voice is inputted from the microphone mentioned above, the sound signal generator (it mentions later in detail) which outputs the sound signal according to that voice to a loudspeaker, and the cell which is the power source of this sound signal generator are built in. 3 is a control unit which consists of two or more push buttons attached in the body section 2, and is constituted by various kinds of carbon buttons shown below.

[0029] (1) It is the carbon button which chooses the mode of operation of the toy in a mode selection carbon button book operation gestalt, and the carbon button is prepared one [ at a time ] corresponding to each following mode, respectively.

**\*\* Learning mode --** The mode in which the toy in this operation gestalt is made to master language or a sound.

**\*\* Setting mode --** The mode in which make it utter among current time, the language made to master in learning mode, or a sound when one of things is inputted into which time of day, what kind of sound, or voice by the microphone, or the correspondence relation is set up.

**\*\* Execute mode --** The mode in which the contents set up in setting mode are performed.

**\*\* Random mode --** The mode in which arbitration is made to utter the language or the sound which memorized the inputted voice to arbitration and the toy in this

operation gestalt has memorized.

**\*\* Message mode --** The mode to which sequential utterance is carried out [ sound / which was memorized / the language or the sound ].

[0030] (2) When setting up time of day in time-of-day setup-key setting mode, carry out a depression (it mentions later in detail).

(3) In current time carbon button setting mode, when making the toy of this operation gestalt notify current time with voice, carry out a depression (it mentions later in detail).

(4) The carbon button which makes the alarm clock function (it mentions later in detail) which the toy in an alarm clock carbon button book operation gestalt has effective or an invalid.

(5) It is constituted by the carbon button corresponding to the figure of ten keys 0-9.

(6) In each mode which carried out decision carbon button \*\*\*\*, in case it opts for a setup etc., it is pushed. Furthermore, 4 is a display and displays the contents set up with the various carbon buttons which constitute the control unit 3 mentioned above.

[0031] Next, the speech recognition mentioned above and the configuration of a speech generation device are explained with reference to drawing 2 . Here, the sign same about the part equivalent to each part shown in drawing 1 is attached,



and the explanation is omitted. In this drawing, 10 is a microphone (henceforth a microphone) and outputs the electrical signal (henceforth a sound signal) according to the voice built in and inputted into the head drawing 1's sewing-based [ 1 ]. 11 is an amplifying circuit and amplifies the sound signal outputted from the microphone 10. 12 is a filter and removes the electrical signal outside an audible band. 13 is an A/D converter, it sampling-izes the sound signal which passed the filter 12 a predetermined period, changes it into a digital signal, and is outputted as voice data.

[0032] 14 is the voice data modification section, and it outputs the voice data stored temporarily with the directions from a control section 29 to either the voice data storage section 15 or D/A converter 16 while it stores temporarily the voice data outputted from A/D converter 13. Moreover, based on the directions from a control section 29, some or all of voice data that is outputted is altered in that case. The voice data storage section 15 stores the voice data outputted from the voice data modification section 14 in the address specified by the control section 29. D/A converter 16 changes into an analog signal the voice data outputted from the voice data modification section 14 or the voice data storage section 15.

[0033] 17 is a voice conversion circuit, and it adjusts suitably the wave of the analog signal outputted from D/A converter 16 so that it may become voice similar to the voice of an actual parakeet. 18 is an amplifying circuit and amplifies

the sound signal outputted from the voice conversion circuit 17. 19 is a sound-volume equalization circuit, and in order to enable accommodation of the sound volume of the voice which the toy in this operation gestalt utters, it adjusts the amplitude of the sound signal outputted from the amplifying circuit 18. 20 is a loudspeaker, is built in a throat sewing-based [ 1 ], and changes and outputs the sound signal outputted from the sound-volume equalization circuit 19 to a sound.

[0034] 21 is the description pattern extract section, asks for linear predictor coefficients, a Fourier-spectrum multiplier, etc. from the sound signal outputted from the filter 12; and extracts the audio description pattern inputted into the microphone 10. 22 is a voice input detecting element, and when the voice inputted into the microphone 10 is more than predetermined level, it outputs a detecting signal to a control section 29. 23 is a timer and clocks current time. 24 is the toy description pattern storage section, and in order that the voice inputted from the microphone 10 may judge whether it is what was emitted from the toy in this another operation gestalt, the description pattern of a sound signal changed by the voice conversion circuit 17 is memorized.

[0035] 25 is the description pattern storage section and memorizes the description pattern outputted from the description pattern extract section 21 to the address to which it was directed by the control section 29. 26 is a counter, and it always carries out counting of the predetermined numerical range after the

power source of the toy in this operation gestalt is turned on until it is turned off. Here, the counter 26 is always counted by numeric-value within the limits of 0-99 (decimal number). 27 is the random-number-generation section and outputs the random number RND by the 3rd place of decimal point by numeric-value within the limits of 0-1 with the directions from a control section 29. 28 is a driving gear and is made to open and close by the motor which does not illustrate beak 1a sewing-based [ 1 ] with the directions from a control section 29. 29 is a control section and controls each part mentioned above.

[0036] Next, the control action of a control section 29 is explained with reference to the flow chart shown in drawing 3 thru/or drawing 8 . First, if the electric power switch which the toy of this operation gestalt does not illustrate is turned ON, a control section 29 will be in the depression standby condition of the mode selection carbon button in a control unit 3. And when a mode selection carbon button is pushed, processing in the mode corresponding to the pushed mode carbon button is started. Hereafter, each processing is explained for every various modes.

[0037] [Processing in learning mode] If the learning mode carbon button of a control unit 3 is pushed, a control section 29 will start processing of the flow chart shown in drawing 3 . Although the voice generated from the toy in this operation gestalt here is the voice fundamentally inputted from the outside, the

toy of this operation gestalt determines whether memorize correctly the voice inputted according to the predetermined probability rather than to only memorize the inputted voice as it is.

[0038] If processing of the flow chart shown in drawing 3 is started, a control section 29 will progress to step Sa1, and will reset the following flags used in learning mode.

- The flag which shows whether it is generated from the toy as the toy in this operation gestalt with the same voice inputted into the toy voice judging flag MAC microphone 10. If generated from the same toy, it will be set to "1", and it will be set to "0" if not generated from the same toy.

- The flag which shows whether the probability to memorize the voice inputted from the reach flag REA outside makes it high. When making high the probability to memorize, it is set to "1", and it is set to "0" when it is still the usual probability.

[0039] In step Sa1, if initialization is completed, it progresses to step Sa2 and a control section 29 will be in the standby condition of the voice inputted from a microphone 10. That is, when it judges whether the detecting signal was outputted from the voice input detecting element 22 and a detecting signal is not outputted, it considers that voice was not inputted from a microphone 10, a decision result serves as NO, and step Sa2 is processed again. And if voice is inputted into a microphone 10 and a detecting signal is outputted from the voice

input detecting element 22, a decision result serves as YES, and when a detecting signal is no longer outputted again, it will progress to processing of step Sa3 after that. Moreover, in the voice data modification section 14, the voice data outputted from A/D converter 13 is stored temporarily in the meantime.

[0040] In step Sa3, a control section 29 compares the description pattern outputted from the description pattern extract section 21 with the description pattern memorized by the toy description pattern storage section 24. And the description pattern which progressed to step Sa4 and was outputted from the description pattern extract section 21, When it is judged that the feature parameter memorized by the toy description pattern storage section 24 is compared, and both are similar The voice inputted into the description pattern microphone 10 regards it as what was emitted from the same toy as the toy of this operation gestalt, and a decision result serves as YES, and it progresses to step Sa6, after progressing to step Sa5 and setting the toy voice judging flag MAC to "1." Moreover, when it is judged that both are not similar, a decision result serves as NO and progresses to step Sa6, without processing step Sa5.

[0041] In step Sa6, a control section 29 reads the counted value CNT of a counter 26. And it judges whether the read value is a constant a. Here, a constant a is any one integer among 0-49, and the value is defined beforehand. And when counted value CNT is in agreement with an integer a, after a decision

result serves as YES, progresses to step Sa8 and sets the reach flag REA to "1", the address with which specific voice data was memorized is specified, and the voice data is made to output to D/A converter 16 to the voice data storage section 15 at step Sa9.

[0042] And after the specific voice data outputted from the voice data storage section 15 is changed into an analog sound signal with D/A converter 16, it is pronounced as specific sounds (for example, chirping peculiar to an actual parakeet etc.) from a loudspeaker 20 through the voice conversion circuit 17, an amplifying circuit 18, and the sound-volume equalization circuit 19. Moreover, a user can recognize that the probability for the toy of this operation gestalt to memorize the inputted voice correctly now became high by hearing this specific sound. And completion of processing of step Sa9 advances a control section 29 to the processing to step Sa10.

[0043] On the other hand, when counted value CNT and an integer a are not in agreement in step Sa7, it progresses to step Sa10, without processing steps Sa8 and Sa9. At step Sa10, the sequential comparison of the description pattern outputted from the description pattern extract section 21 and the description pattern memorized from address [ of the description pattern storage section 25 ] P (0) by P (i) is carried out, and it judges whether there is any description pattern which is in agreement with the description pattern storage section 25.

[0044] And when there is no description pattern which progresses to step Sa11 and is in agreement, it is set to NO and progresses to step Sa12, and a decision result increments the value of Variable i one time, makes address [ of the description pattern storage section 25 ] P (i) memorize the description pattern outputted from the description pattern extract section 21 at step Sa13, and progresses to step Sa15. On the other hand, when there is a pattern which is in agreement in the description pattern storage section 25 in step Sa11, a decision result serves as YES, it progresses to step Sa14, and the value of i of address P (i) the pattern is remembered to be is detected.

[0045] Next, it progresses to step Sa15, and in order that the voice inputted into the microphone 10 this time may count how many times it was inputted until now, one value of the count count data C of an input (i) is incremented. Subsequently, it progresses to step Sa16 and judges whether the count of an input of the voice inputted into the microphone 10 this time became 30 times or more. Here, when the count of an input is less than 30 times, a decision result serves as NO and progresses to step Sa17.

[0046] Subsequently, at step Sa17, it judges whether either the toy voice judging flag MAC or the reach flag REA is set to "1." Here, when neither the toy voice judging flag MAC nor the reach flag REA is set to "1", a decision result serves as NO and progresses to step Sa18. And a control section 29 makes a random

number RND output to the random-number-generation section 27, and generates an integral random number in numeric-value within the limits of 0-29 based on the random number RND.

[0047] Subsequently, it progresses to step Sa19 and the random number obtained by processing in step Sa18 judges whether it is in agreement with a constant b. Here, a constant b is any one integer among 0-29, and the value is defined beforehand. And when the random number obtained at step Sa18 and a constant b are not in agreement, a decision result serves as NO, and it progresses to step Sa20, and directs to alter some or all of voice data that was inputted this time which is stored temporarily to the voice data modification section 14. Thereby, the voice data modification section 14 changes some or all of the above-mentioned voice data into intentionally different data.

[0048] Next, a control section 29 progresses to processing of step Sa21, and makes the voice data stored temporarily to the voice data modification section 14 output to D/A converter 16. And after the voice data outputted to D/A converter 16 is changed into an analog sound signal, it is changed into a tone like a parakeet in the voice conversion circuit 17, and voice is uttered from a loudspeaker 20 through an amplifying circuit 18 and the sound-volume equalization circuit 19. Moreover, the contents of the voice uttered differ from the contents of the voice inputted into the microphone 10 at this time. Here, in this



operation gestalt, when a detecting signal is outputted from the voice input detecting element 22 while voice was uttered from the loudspeaker 20 (i.e., when a certain sound is inputted into a microphone 10), a control section 29 shall consider that the sound is an invalid.

[0049] On the other hand, when the random number obtained at step Sa18 and a constant b are in agreement in step Sa19, processing in step Sa20 is not performed, but it progresses to processing of the direct step Sa21. the contents of the voice inputted from the microphone 10 by this -- it remains as it is and, moreover, is emitted from a loudspeaker 20 in a tone like a parakeet.

[0050] In addition, when the count of an input of the voice inputted into the microphone 10 this time is judged to be 30 times or more in step Sa16, Or when it is judged in step Sa17 that either the toy voice judging flag MAC or the reach flag REA is set to "1" Progressing to step Sa22, a control section 29 makes a random number RND output to the random-number-generation section 27, and generates an integral random number in numeric-value within the limits of 0-2 based on the random number RND.

[0051] And it progresses to step Sa23 and the random number obtained by processing in step Sa22 judges whether it is in agreement with a constant c. Here, a constant c is any one integer among 0-2, and the value is defined beforehand. And when the random number obtained at step Sa22 and a

constant c are not in agreement, a decision result serves as NO and it progresses to step Sa20, and when in agreement, it progresses to step Sa21 and the same processing as the processing mentioned above is performed hereafter.

[0052] Subsequently, after processing of step Sa21 is completed, a control section 29 progresses to step Sa24, and directs to drive beak 1a sewing-based [ 1 ] to a driving gear 27. Thereby, a driving gear 27 makes beak 1a sewing-based [ 1 ] open and close by driving the motor which is not illustrated, while the sound has occurred from the loudspeaker 20. Thereby, sewing-basis [ 1 ] is visible as it has talked.

[0053] Next, a control section 29 progresses to step Sa25, and judges whether the decision carbon button of a control unit 3 was pushed. Here, a user hears the sound emitted from the loudspeaker 20, and if the depression of the decision carbon button of a control unit 3 is carried out when the same language as the language which spoke to the microphone 10 is uttered, the decision result of step Sa25 will serve as YES, and will progress to step Sa26. Moreover, when a predetermined time decision carbon button is not pushed, return and the processing mentioned above are again repeated to step Sa2. If a decision carbon button is pushed by the user and a control section 29 progresses to processing of step Sa26, a control section 29 will make address [ of the voice

data storage section 15 ] V (i) memorize the voice data stored temporarily in the voice data modification section 14, and will end processing of learning mode.

[0054] Since the voice inputted into the microphone 10 is correctly uttered only at 1 time of a rate to 30 times in learning mode as explained above, rather than what memorizes the inputted voice as it is simply, parakeet-likeness can be directed and, moreover, unexpected nature can be given. Moreover, since the rate uttered correctly increases when the same voice is inputted more than the count of predetermined, the phenomenon in which it does not utter as it inputted very much, even if it inputted the voice same how many times decreases.

[0055] Furthermore, a hope can be given by sensibility like [ telling a user about the rate which is made to increase the rate pronounced correctly very rarely intentionally, and is pronounced correctly having increased / a user ] the reach sound at the time of a pachinko game. Moreover, way another again of enjoying oneself becomes possible, since the rate pronounced correctly increases when the voice inputted from the microphone 10 is uttered from the same toy as the toy of this operation gestalt, for example, when the toy concerned is gathered, distinction-ization of mastering voice easily is made when each other voice is learned and it suits with the same toys, and a user makes his voice master.

[0056] [Processing in setting mode] If the setting mode carbon button of a control unit 3 is pushed, a control section 29 will start processing of the flow chart shown

in drawing 4 and drawing 5 . Although setting mode is the mode in which it sets up whether current time or the voice made to memorize with the learning mode mentioned above is made to utter to what kind of case, there are the following five kinds of the contents of the setup.

- Make the memorized specific voice utter to the specific language or the specific sound which setting 1 user emitted. for example, the user emitted -- "-- now -- " -- \*\* -- the word "return" is made to utter [ which it has memorized ] to the language to say

[0057] - Make voice notify current time to the specific language or the specific sound which setting 2 user emitted. For example, the voice "it is a part (current time) for \*\*\*\* at the time of OO" is made to utter to the word which the user emitted of "being ? when."

- Setting 3 user makes the time of day set up with the ten key of a control unit 3 notify current time with voice.

[0058] - Setup 4 (alarm clock function)

If a user makes the specific voice specified beforehand repeat and utter at the time of day set up with the ten key of a control unit 3 and a user's specific words are inputted, utterance of the voice will be stopped. For example, if the specified time of day comes, the words "good morning" will be repeated and uttered and utterance will be stopped with a user's word of "having occurred."

- Setting 5 user makes the specific voice beforehand specified as the time of day set up with the ten key of a control unit 3 utter. For example, when the specified time of day comes, the word "it is the time amount of television" is made to utter.

[0059] Processing of the control section 29 when performing hereafter various setup mentioned above is explained with reference to drawing 4 and drawing 5.

First, if a user does the depression of the setting mode carbon button of a control unit 3, a control section 29 will progress to step Sb1, and will judge whether voice was inputted from the microphone 10. And when a detecting signal is not outputted from the sound signal detecting element 22, a decision result serves as NO and progresses to step Sb2. At step Sb2, it judges whether the time-of-day setup key of a control unit 3 was pushed. Here, when the time-of-day setup key of a control unit 3 is not pushed, a decision result serves as NO and returns to processing of step Sb1. Thus, first, a control section 29 repeats processing of step Sb1 and step Sb2 until voice is inputted or a time-of-day setup key is pushed.

[0060] In this condition, if voice is inputted into a microphone 10, a detecting signal will be outputted from the voice input detecting element 22, and, thereby, a control section 29 will recognize that voice was inputted. And when a detecting signal is no longer outputted from the voice input detecting element 22, it progresses to step Sb3. And in step Sb3, when a control section 29 judges

whether the ten key of a control unit 3 was pushed and a ten key is not pushed, a decision result serves as NO and progresses to step Sb4. At step Sb4, it judges whether the current time carbon button of a control unit 3 was pushed. Here, when a current time carbon button is not pushed, a decision result serves as NO and returns to processing of step Sb3. Thus, a control section 29 repeats processing of step Sb3 and step Sb4 until a ten key is pushed or a current time carbon button is pushed.

[0061] And if it detects that the ten key was pushed for example, a control section 29 will progress to processing of step Sb5, and will perform processing about setup 1. First, after memorizing the input value by the ten key as ten key data n, it progresses to step Sb6 and the voice data memorized by address [ of the voice data storage section 15 ] V (n) is made to output to D/A converter 16 in step Sb5. The voice data memorized by this to the address specified by the ten key is uttered from a loudspeaker 20.

[0062] Next, it progresses to step Sb7 and judges whether the decision carbon button was pushed. Here, when a decision carbon button is not pushed, a decision result serves as NO and returns to step Sb3. And when a ten key is pushed again, it progresses to step Sb5 and the voice data memorized to the address specified by the ten key is made to pronounce from a loudspeaker 20. Thus, a user chooses the voice made to pronounce to the voice inputted into the

microphone 10 with a ten key.

[0063] And when the voice for which it asks is uttered from a loudspeaker 20, a user does the depression of the decision carbon button. Then, the decision result in step Sb7 serves as YES, it progresses to step Sb8, and correspondence with the voice inputted into the microphone 10 and the voice which the toy of this operation gestalt utters to the voice is set up. That is, in step Sb8, a control section 29 is memorized as the address RA 1 which reads selected address V (n) in case setup 1 is performed, i.e., the read-out address. Subsequently, it progresses to step Sb9, and in case setup 1 is performed for the description pattern outputted from the description pattern extract section 21, the address PS 1 of the description pattern storage section 25 is made to memorize as a description pattern for judging the voice which the user uttered.

[0064] Then, a control section 29 progresses to step Sb10, sets to "1" the setting 1 flag EST 1 which shows that setup 1 was made by the user, and progresses to step Sb11. At step Sb11, a user judges whether all setup was completed in setting mode. That is, when the decision carbon button of a control unit 3 is pushed, a decision result serves as YES and the processing in setting mode is ended.

[0065] Moreover, when a decision carbon button is not pushed, a decision result serves as NO, and it progresses to step Sb12, and judges whether carbon

buttons other than a decision carbon button were pushed. Here, when carbon buttons other than a decision carbon button are pushed, a decision result serves as YES, and it is regarded as that to which a user performs other setup succeedingly, and returns to processing of step Sb1. Moreover, when a carbon button is not pushed at all, a decision result serves as NO, it returns to step Sb11, and hereafter, processing of steps Sb11 and Sb12 is repeated until a decision carbon button or the other carbon button is pushed.

[0066] Next, when a ten key is not pushed in step Sb3 and the current time carbon button of a control unit 3 is pushed in step Sb4, the decision result in step Sb4 serves as YES, and a control section 29 performs processing about setup 2. That is, it progresses to step Sb13, and in case setup 2 is performed for the description pattern outputted from the description pattern extract section 21, the address PS 2 of the description pattern storage section 25 is made to memorize as a description pattern for judging the voice which the user uttered.

[0067] And it progresses to step Sb14, and the setting 2 flag EST 2 which shows that setup 2 was made by the user is set to "1", and it progresses to step Sb11. And when a user judges whether all setup was completed in setting mode and a decision carbon button is pushed, it considers that all setup was completed and the processing in setting mode is ended.

[0068] Now, when it is judged that a sound signal next is not inputted in step Sb1



and the time-of-day setup key of a control unit 3 is pushed in step Sb2, the decision result in step Sb2 serves as YES, and a control section 29 performs setup 3 thru/or processing about either of 5. The detail is explained below. First, if the decision result in step Sb2 serves as YES, it will progress to step Sb15, and will consider that the value inputted from the ten key of a control unit 3 is the time of day specified by the user, and the value will be memorized as appointed time-of-day data T.

[0069] Next, it progresses to step Sb16 and judges whether the current time carbon button of a control unit 3 was pushed. Here, if a current time carbon button is pushed, a decision result will serve as YES, it will progress to step Sb17, and a control section 29 will perform processing about setup 3. That is, the appointed time-of-day data T mentioned above are memorized as appointed time-of-day T3 in setup 3. And it progresses to step Sb18, and the setting 3 flag EST 3 which shows that setup 3 was made by the user is set to "1", and it progresses to step Sb11. Then, as mentioned above, when it judges whether all setup [ in / in a user / setting mode ] was completed and a decision carbon button is pushed, it considers that all setup was completed and the processing in setting mode is ended.

[0070] Moreover, in step Sb16, when a current time carbon button is not pushed, a decision result serves as NO and progresses to step Sb19. At step Sb19, it

judges whether voice was inputted from the microphone 10. Here, when a detecting signal is outputted from the voice input detecting element 22, a decision result serves as YES noting that voice is inputted. And it progresses to step Sb20 and a control section 29 performs processing about setup 4 hereafter. First, in case a control section 29 performs setup 4 for the description pattern outputted from the description pattern extract section 21, the address PS 4 of the description pattern storage section 25 is made to memorize it in step Sb20 as a description pattern for judging the voice which the user uttered.

[0071] Next, it progresses to step Sb21 and a control section 29 will be in the depression standby condition of a ten key. That is, if a decision result serves as NO, processing of step Sb21 is repeated and a ten key is pushed until either of the ten keys is pushed, a decision result will serve as YES and will progress to step Sb22. And after substituting the input value of a ten key for the ten key data n, it progresses to step Sb23 and the voice data memorized by address [ of the voice data storage section 15 ] V (n) is made to output to D/A converter 16 at step Sb22. The voice data memorized by this to the address specified by the ten key is uttered from a loudspeaker 20.

[0072] Next, it progresses to step Sb24 and judges whether the decision carbon button was pushed. Here, when a decision carbon button is not pushed, it is set to NO, and a decision result returns to step Sb21, and will be in the depression

standby condition of a ten key again. Hereafter, a control section 29 repeats processing of steps Sb21-Sb24 until a decision carbon button is pushed. Thereby, a user can choose the voice of the request made to utter in case setup 4 is performed with a ten key with a ten key.

[0073] And if the voice for which a user asks from a loudspeaker 20 is uttered and a user does the depression of the decision carbon button, the decision result in step Sb24 will serve as YES, and a control section 29 will perform processing after step Sb25. First, in step Sb25, a control section 29 is memorized as the address RA 4 which reads address V (n) mentioned above in case setup 4 is performed, i.e., the read-out address. Subsequently, it progresses to step Sb26 and the time-of-day data T specified by the user in step Sb15 are memorized as appointed time-of-day T four in setup 4.

[0074] Then, a control section 29 progresses to step Sb27, sets to "1" the setting 4 flag EST 4 which shows that setup 4 was made by the user, and progresses to step Sb11. At step Sb11, a user judges whether all setup was completed in setting mode. That is, when the decision carbon button of a control unit 3 is pushed, a decision result serves as YES, a user considers that all setup was completed and the processing in setting mode is ended.

[0075] Next, in step Sb19, when it is judged that voice was not inputted into a microphone 10, a control section 29 progresses to step Sb28, and judges

whether the ten key was pushed. And when a ten key is pushed, a decision result serves as YES, and processing about the setup 5 after step Sb29 is performed. In addition, in step Sb28, when a ten key is not pushed, a decision result serves as NO and returns to step Sb16. By this, the setting mode carbon button of a control unit 3 is pushed, and first, a control section 29 will process setup 3 thru/or either of 5 according to either the input of the voice to the depression of a current time carbon button, and the toy of this operation gestalt, or the depression of a ten key, after a time-of-day setup key is pushed and time of day is set up from a ten key.

[0076] When a control section 29 progresses to processing of step Sb29, it progresses to step Sb30 and makes the voice data memorized by address [ of the voice data storage section 15 ] V (n) output to D/A converter 16, after a control section 29 substitutes the input value of a ten key for Variable n. The voice data memorized by this to the address specified by the ten key is uttered from a loudspeaker 20. Next, a control section 29 progresses to step Sb31, and judges whether the decision carbon button was pushed. Here, when a decision carbon button is not pushed, a decision result serves as NO and returns to step Sb16. Moreover, if a ten key is pushed while passing through processing of step Sb16 ->Sb19 ->Sb28, the decision result of step Sb28 will serve as YES again, and processing of step Sb29 will be performed.

[0077] And if the voice for which a user asks is uttered and a user does the depression of the decision carbon button from a loudspeaker 20, the decision result in step Sb31 will serve as YES, and a control section 29 will perform processing after step Sb32. Here, in the processing after step Sb32, correspondence with the time of day specified by the user and the voice which the toy of this operation gestalt utters in the time of day is set up. First, in step Sb32, a control section 29 is memorized as the address RA 5 which reads address V (n) mentioned above in case setup 5 is performed, i.e., the read-out address. Subsequently, it progresses to step Sb33 and the time-of-day data T specified by the user in step Sb15 are memorized as appointed time of day T5 in setup 5.

[0078] Then, a control section 29 progresses to step Sb34, sets to "1" the setting 5 flag EST 5 which shows that setup 5 was made by the user, and progresses to step Sb11. At step Sb11, a user judges whether all setup was completed in setting mode. That is, when the decision carbon button of a control unit 3 is pushed, a decision result serves as YES, a user considers that all setup was completed and the processing in setting mode is ended.

[0079] After the setting mode carbon button of a control unit 3 is pushed in setting mode, voice is first inputted into a microphone 10, when a ten key is pushed succeedingly, processing of setup 1 is performed, and as explained

above, when a current time carbon button is pushed, processing of setup 2 is performed. moreover, the case where processing of setup 3 is succeedingly inputted into voice when a current time carbon button is pushed when a setting mode carbon button is pushed, the introduction time-of-day setup key is pushed first and a time-of-day setup is subsequently made -- processing of setup 4 -- moreover, setup 5 is processed when a ten key is pushed.

[0080] [Processing in execute mode] If the execute mode carbon button of a control unit 3 is pushed, a control section 29 will start processing of the flow chart shown in drawing 6 and drawing 7 . Execute mode is the mode in which the contents set up in the setting mode mentioned above are performed, and explains actuation of the control section 29 in execute mode hereafter. If an execute mode carbon button is pushed, a control section 29 will process the step Sc 1 in the flow chart of drawing 6 (a) first. That is, the condition of each setting flags EST1-EST5 is searched, and the flag with which "1" is set is detected.

[0081] And progress to a step Sc 2, the information on current time is made to output to a timer 23, and the information is memorized as current time data Tc. And it progresses to a step Sc 3 and the contents of a setting are performed by a unit of 1 time about the setting flag with which "1" is set based on the result searched at a step Sc 1, respectively. In addition, about the processing in a step Sc 3, it mentions later. Next, it progresses to a step Sc 4 and judges whether the

random mode carbon button of a control unit 3 is pushed.

[0082] And when the random mode carbon button is pushed, a decision result serves as YES, it progresses to a step Sc 5, and processing of a random mode is performed. In addition, this processing is explained in detail later. Moreover, when the random mode carbon button is not pushed, a decision result is set to NO, and it progresses to a step Sc 6, without processing a step Sc 5, and judges whether other mode carbon buttons were pushed. Here, when other mode carbon buttons are pushed, while a decision result serves as YES and ends processing of execute mode, the mode of the pushed mode carbon button is performed. Moreover, in a step Sc 6, when other mode carbon buttons are not pushed, a decision result serves as NO, and processing of steps Sc2-Sc6 is repeated and performed return and the following to a step Sc 2.

[0083] Next, in the step Sc 3 mentioned above, the processing actuation at the time of performing each setup is explained with reference to drawing 6 and drawing 7 . First, when the setting 1 flag EST 1 is set to "1", a control section 29 performs processing of the flow chart shown in drawing 6 (b). In a step Sc 7, a control section 29 judges whether voice was inputted into the microphone 10. Here, a detecting signal is outputted from the voice input detecting element 22, and if it judges that voice was inputted, a decision result will serve as YES and will process a step Sc 8. Moreover, when a detecting signal is not outputted from

the voice input detecting element 22, a decision result serves as NO and executive operation about other setting flags set to the step Sc 3 of drawing 6 (a) by return and "1" is performed.

[0084] And if the decision result in a step Sc 7 serves as YES and progresses to a step Sc 8, a control section 29 compares with the description pattern outputted from the description pattern extract section 21, and the description pattern memorized to the address PS 1 of the description pattern storage section 25. And it progresses to a step Sc 9 and judges whether both description pattern was in agreement. When both description pattern is not in agreement, a decision result serves as NO, it returns to the step Sc 3 of drawing 6 (a), and executive operation about other setting flags set to "1" is performed. Moreover, when in agreement, it is directed that the voice data memorized to the address of the voice data storage section 15 which a decision result serves as YES, progresses to a step Sc 10, and is specified by the read-out address RA 1 outputs to D/A converter 16.

[0085] And after the voice data outputted to D/A converter 16 is changed into an analog sound signal and changed into a tone like a parakeet in the voice conversion circuit 17, voice is uttered from a loudspeaker 20 through an amplifying circuit 18 and the sound-volume equalization circuit 19. The description pattern equivalent to the voice to say is memorized. thereby -- for



example, the address PS 1 of the description pattern storage section 25 -- "-- now -- " -- \*\* -- the case where the voice data "return" is memorized to the address of the voice data storage section 15 specified by the read-out address RA 1 -- a user -- a microphone 10 -- "-- now -- " -- \*\*, if the voice to say is inputted The voice "return" is uttered in a tone like a parakeet from a loudspeaker 20.

[0086] Moreover, it is visible as beak 1a sewing-based [ 1 ] opened and closed beak 1a sewing-based [ 1 ] while the sound had occurred from the loudspeaker 20 by this, and it points to a driving gear 27 at this time so that a control section 29 may be driven, and sewing-basis [ 1 ] has spoken it. And after finishing processing of a step Sc 10, a control section 29 performs executive operation of the setup about return and other setting flags set to "1" to the step Sc 3 of drawing 6 (a).

[0087] Next, processing actuation of the control section 29 when the setting 2 flag EST 2 is set to "1" is explained with reference to the flow chart shown in drawing 6 (c). First, a control section 29 progresses to a step Sc 11, and judges whether voice was inputted into the microphone 10. Here, a detecting signal is outputted from the voice input detecting element 22, and if it judges that voice was inputted, a decision result will serve as YES and will process a step Sc 12. Moreover, when a detecting signal is not outputted from the voice input detecting element 22, a decision result serves as NO and executive operation of the setup

is performed about other setting flags set to the step Sc 3 of drawing 6 (a) by return and "1."

[0088] And if the decision result in a step Sc 11 serves as YES and progresses to a step Sc 12, a control section 29 compares with the description pattern outputted from the description pattern extract section 21, and the description pattern memorized to the address PS 2 of the description pattern storage section 25. And it progresses to a step Sc 13 and judges whether both description pattern was in agreement. When both description pattern is not in agreement, a decision result serves as NO, it returns to the step Sc 3 of drawing 6 (a), and executive operation of the setup is performed about other setting flags set to "1." Moreover, when in agreement, a decision result serves as YES, and it progresses to a step Sc 14, and directs to output the voice data corresponding to the current time data Tc to D/A converter 16 to the voice data storage section 15.

[0089] And after the voice data corresponding to the current time outputted to D/A converter 16 is changed into an analog sound signal and changed into a tone like a parakeet in the voice conversion circuit 17, current time is emitted from a loudspeaker 20 through an amplifying circuit 18 and the sound-volume equalization circuit 19. If a user inputs into a microphone 10 the voice of "being ? when" when the description pattern which is equivalent to the address PS 2 of the description pattern storage section 25 at the voice of "being ? when" is

memorized by this, from a loudspeaker 20, current time will be emitted in a tone like a parakeet.

[0090] And after finishing processing of a step Sc 14, a control section 29 performs executive operation of the setup about other setting flags set to the step Sc 3 of drawing 6 (a) by return and "1."

[0091] Next, processing actuation of the control section 29 when the setting 3 flag EST 3 is set to "1" is explained with reference to the flow chart shown in drawing 7 (a). First, a control section 29 progresses to a step Sc 15, compares the current time data Tc with appointed time-of-day data T3 set up by the user by the setup 3 in setting mode, and judges whether it is in agreement. And when both are in agreement, a decision result serves as YES, it progresses to a step Sc 16, and it is directed to the voice data storage section 15 that the voice data corresponding to the current time data Tc outputs to D/A converter 16.

[0092] And after the voice data corresponding to current time outputted to D/A converter 16 is changed into an analog sound signal and changed into a tone like a parakeet in the voice conversion circuit 17, current time is emitted from a loudspeaker 20 through an amplifying circuit 18 and the sound-volume equalization circuit 19. Thereby, if the toy in this operation gestalt becomes the time of day specified by the setup 3 in setting mode, it will utter current time in voice like a parakeet. And after finishing processing of a step Sc 16, a control

section 29 performs executive operation of the setup about other setting flags set to the step Sc 3 of drawing 6 (a) by return and "1." On the other hand, when appointed time-of-day data T3 and the current time data Tc are not in agreement in a step Sc 15, a decision result serves as NO and it returns to the step Sc 3 of drawing 6 (a), without processing a step Sc 16.

[0093] Next, processing actuation of the control section 29 when the setting 4 flag EST 4 is set to "1" is explained with reference to the flow chart shown in drawing 7 (b). First, a control section 29 progresses to a step Sc 17, and judges whether the alarm clock carbon button of a control unit 3 is pushed. And when the alarm clock carbon button is not pushed, a decision result serves as NO and executive operation of the setup is performed about other setting flags set to the step Sc 3 of drawing 6 (a) by return and "1."

[0094] Moreover, when the alarm clock carbon button is pushed, a decision result serves as YES and progresses to processing of a step Sc 18. And in a step Sc 18, a control section 29 is the setup 4 in setting mode, compares the current time data Tc with appointed time-of-day data T four set up by the user, and judges whether it is in agreement. And when a decision result is set to YES and it progresses to a step Sc 19, when both are in agreement, and both are not in agreement, a decision result serves as NO and returns to the step Sc 3 of drawing 6 (a).

[0095] If it progresses to a step Sc 19, a control section 29 directs that the voice data memorized to the address of the voice data storage section 15 specified by the read-out address RA 4 outputs to D/A converter 16. And after the voice data outputted to D/A converter 16 is changed into an analog sound signal and changed into a tone like a parakeet in the voice conversion circuit 17, voice is uttered from a loudspeaker 20 through an amplifying circuit 18 and the sound-volume equalization circuit 19.

[0096] Subsequently, a control section 29 progresses to a step Sc 20, and judges whether voice was inputted into the microphone 10. Here, a detecting signal is outputted from the voice input detecting element 22, and if it judges that voice was inputted, a decision result will serve as YES and will process a step Sc 21. Moreover, when a detecting signal is not outputted from the voice input detecting element 22, a decision result serves as NO, it returns to a step Sc 19 again, and processing for making the voice data memorized to the address of the voice data storage section 15 utter specified by the read-out address RA 4 is performed.

[0097] And if the decision result in a step Sc 20 serves as YES and progresses to a step Sc 22, a control section 29 compares with the description pattern outputted from the description pattern extract section 21, and the description pattern memorized to the address PS 4 of the description pattern storage section

25. And it progresses to a step Sc 22 and judges whether both description pattern was in agreement. When both description pattern is not in agreement, a decision result serves as NO, it returns to a step Sc 19, and processing for making it utter based on the voice data memorized to the address of the voice data storage section 15 again specified by the read-out address RA 4 is performed.

[0098] Moreover, if the description pattern outputted from the description pattern extract section 21 and the description pattern memorized to the address PS 4 of the description pattern storage section 25 are in agreement, a decision result will serve as YES and will return to the step Sc 3 of drawing 6 (a). Thus, in the flow chart of drawing 7 (b), if the time of day which the alarm clock carbon button of a control unit 3 is pushed, and was specified by the setup 4 in setting mode comes, the voice specified beforehand will be repeated and uttered from a loudspeaker 20 until a user utters predetermined voice.

[0099] Next, processing actuation of the control section 29 when the setting 5 flag EST 5 is set to "1" is explained with reference to the flow chart shown in drawing 7 (c). First, a control section 29 judges whether by the setup 5 in setting mode, the appointed time-of-day data T5 set up by the user are compared with the current time data Tc, and it is in agreement in a step Sc 23. And when a decision result is set to YES and it progresses to a step Sc 24, when both are in

agreement, and both are not in agreement, a decision result serves as NO and returns to the step Sc 3 of drawing 6 (a).

[0100] If it progresses to a step Sc 24, a control section 29 directs that the voice data memorized to the address of the voice data storage section 15 specified by the read-out address RA 5 outputs to D/A converter 16. And after the voice data outputted to D/A converter 16 is changed into an analog sound signal and changed into a tone like a parakeet in the voice conversion circuit 17, voice is uttered from a loudspeaker 20 through an amplifying circuit 18 and the sound-volume equalization circuit 19. Then, a control section 29 returns to the step Sc 3 of drawing 6 (a). Thus, in the flow chart of drawing 7 (c), if the time of day specified by the setup 5 in setting mode comes, the voice specified by the user will be uttered.

[0101] In addition, although it will become adiaborous [ the toy in this operation gestalt ] if the voice which was not matched at the time of setting mode is inputted, in such a case 29, for example, a control section, you choose the voice data of arbitration from the voice data memorized by the voice data storage section 15, and may make it make it output to D/A converter 16 in the execute mode mentioned above.

[0102] [Processing in a random mode] Processing of a random mode is made during operation of execute mode like [ it is \*\*\*\*\* and ] also from the flow chart of

drawing 6 (a). Hereafter, the processing in a random mode is explained with reference to the flow chart of drawing 8 . First, in the step Sc 4 of the flow chart of drawing 6 (a), when it is judged that the random mode carbon button of a control unit 3 is pushed, a control section 29 progresses to a step Sc 5, and starts the processing in a random mode. That is, first, it progresses to step Sd1 in the flow chart of drawing 8 , and judges whether voice was inputted to the microphone 10.

[0103] Here, a detecting signal is outputted from the voice input detecting element 22, and when it is judged that voice was inputted to the microphone 10, it is determined based on a random number whether memorize the voice inputted by a decision result serving as YES. That is, it progresses to step Sd2 and a random number RND is made to output to the random-number-generation section 27. Next, it progresses to step Sd3 and the random number RND and constant d which were outputted from the random-number-generation section 27 are compared. Here, a constant d is a predetermined decimal displayed by numeric-value within the limits of 0-1 till the 3rd place of decimal point. And when a random number RND and a constant d are in agreement, a decision result serves as YES, after making the address VR of the voice data storage section 15 (j) memorize the audio voice data progressed and inputted into step Sd4, it progresses to step Sd5 and the value of j is incremented one time. Here, the address VR mentioned above means the memory area where the voice data



memorized at the time of a random mode is memorized in the voice data storage section 15.

[0104] And when it is judged that a sound signal was not inputted at step Sd1 (a decision result is NO), when a random number RND and a constant d are not in agreement at step Sd3 (a decision result is NO) When processing of step Sd5 is completed, it progresses to step Sd6 and determines whether generate one which is memorized by the voice data storage section 15 of voice data based on a random number. That is, the random number RND which was made to output a random number RND to the random-number-generation section 27 in step Sd6, and was subsequently outputted from the random-number-generation section 27 at step Sd7 judges whether it is in agreement with a constant e. Here, a constant e is a predetermined decimal which was displayed by numeric-value within the limits of 0-1 till the 3rd place of decimal point and from which a constant d differs.

[0105] And when a random number RND and a constant e are not in agreement, a decision result serves as NO and it returns to the step Sc 6 in the flow chart of drawing 6 (a). Moreover, when a random number RND and a constant e are in agreement, a decision result serves as YES, and it progresses to step Sd8, and it is determined which voice data is uttered among the voice data memorized by the voice data storage section 15 based on a random number. That is, first, a random number RND is made to output to the random-number-generation

section 27, and the random number of numeric-value within the limits of 0 - j is computed based on the variable j shown at this random number RND and step Sd5.

[0106] Next, it progresses to step Sd9 and voice data is made to output to the voice data storage section 15 from the address VR (k) specified with the random number obtained at step Sd8. Thereby, after the voice data outputted from the voice data storage section 15 is changed into an analog sound signal by D/A converter 16 and being changed into a tone like a parakeet in the voice conversion circuit 17, voice is uttered from a loudspeaker 20 through an amplifying circuit 18 and the sound-volume equalization circuit 19. And after finishing processing of step Sd9, a control section 29 progresses to the step Sc 6 of drawing 6 (a), and continues the processing in execute mode.

[0107] Thus, when a random mode is chosen by the user, whenever it passes the step Sc 5 of the flow chart of drawing 6 (a), processing of the flow chart shown in drawing 8 is made. Thereby, the toy in this operation gestalt incorporates to arbitration the voice inputted into the microphone 10, and utters the language of arbitration to arbitration.

[0108] [Processing in message mode] If the message mode carbon button of a control unit 3 is pushed, a control section 29 will start processing of the flow chart shown in drawing 9 . First, a control section 29 clears the value of the

variable m for addressing of the voice data storage section 15 in message mode in a step Se 1. And it progresses to a step Se 2 and the voice data memorized by address V (m) is made to output to D/A converter 16 to the voice data storage section 15. Thereby, after the voice data outputted to D/A converter 16 is changed into an analog sound signal and changed into a tone like a parakeet in the voice conversion circuit 17, voice is uttered from a loudspeaker 20 through an amplifying circuit 18 and the sound-volume equalization circuit 19.

[0109] Next, it progresses to a step Se 3, the value of the variable m for addressing is incremented one time, and it judges whether the value of m exceeded the variable i for addressing in learning mode in a step Se 4. And when the value of m is below a value of Variable i, a decision result serves as NO, it returns to a step Se 2, and the voice data memorized to the next address of the voice data storage section 15 is made to output to D/A converter 16. Thus, the voice data which made increase every one value of m and the toy of this operation gestalt was made to master at the time of learning mode is made to utter from a loudspeaker 20 one by one. And in a step Se 4, when it is judged that the value of m exceeded the value of Variable i, a decision result serves as YES and the processing in message mode is ended.

[0110] Thus, in message mode, since the voice data which the voice data storage section 15 was made to memorize at the time of learning mode is made

to utter altogether one by one, the contents of the voice data which the toy of this operation gestalt has memorized can be checked. Moreover, since the message made to master can be heard when the partner pushes a message mode carbon button after that if the toy of this operation gestalt is made to master a message to tell a partner at the time of learning mode for example, it can use also as a messenger cable who changes the toy of this operation gestalt into a tone like a parakeet.

[0111] As explained above, according to the toy in this operation gestalt, the voice which answers can be set as arbitration to the voice and the sound which were inputted. Thereby, unexpected nature can be given to the reaction of a toy and the toy which cannot get bored easily can be offered.

[0112] In addition, in the operation gestalt mentioned above, although sewing-basis which imitated the parakeet was used, sewing-basis of other animals, such as a dog besides an animal, and a cat, which emits not only a parakeet but \*\*\*\* to produce, such as a hill myna, and when it may use a doll etc. and allowances are in these dolls or the sewing-based volume, you may also incorporate the speech recognition and the speech generation device which were mentioned above to those interior. Moreover, you may also build into an indoor furniture etc. the speech recognition and the speech generation device which were mentioned above.

[0113] Moreover, although he is trying to generate the voice data which adjusted the wave of the analog signal outputted from D/A converter 16, and was read from the voice data storage section 15 by the voice conversion circuit 17 in voice like a parakeet, it constitutes so that it can change into two or more kinds of tone quality, and you may enable it to choose the tone quality which forms and changes a selecting switch etc. in the voice conversion circuit 17.

[0114] Furthermore, the function of the toy mentioned above may be realized on a personal computer (henceforth a personal computer), or a video game machine. In that case, when displaying images, such as a parakeet, on monitor display and making it generate voice from a personal computer as a substitute sewing-based [ 1 ]; it may be made to perform an image processing to which a beak, feather, etc. of a parakeet on a screen are moved. Here, the configuration in the case of realizing the function of the toy which used and mentioned the personal computer above to drawing 10 is shown.

[0115] For a microphone and 20, as for the body of a personal computer, and 51, in drawing 10 , a loudspeaker and 50 are [ 10 / a keyboard and 52 ] monitors. Moreover, 53 is a sound-source board built into the body 50 of a personal computer, analog-izes the digital data from a bus, and outputs it to a loudspeaker 20 while providing an amplifying circuit, a filter, an A/D converter, a D/A converter, etc., digitizing the sound signal outputted from a microphone 10

and outputting to the bus within the body 50 of a personal computer.

[0116] In the configuration shown in drawing 10 , the amplifying circuits 11 and 18 shown in drawing 2 ; A/D converter 13, and D/A converter 16 correspond to the sound-source board 53, and the control unit 3 of drawing 2 corresponds to input devices, such as a keyboard 52 or a mouse (illustration abbreviation). Moreover, the voice data storage section 15 in drawing 2 , the toy description pattern storage section 24, and the description pattern storage section 25 are realizable by using RAM within the body 50 of a personal computer. Furthermore, processing in the voice data modification section 14 of drawing 2 , the description pattern transducer 21, and the voice input detecting element 22 is performed by the body 50 of a personal computer based on the digitized sound signal which is outputted from the sound-source boat 53. Moreover, the function of the timer 23 of drawing 2 , a counter 26, and the random-number-generation section 27 is realizable by using the internal clock which the body 50 of a personal computer possesses, RAM, etc.

[0117] And the program for making the body 50 of a personal computer perform procedure shown in the flow chart of drawing 3 thru/or drawing 9 is stored in record media, such as a floppy disk, and the configuration which shows the function of a toy in which the above-mentioned program was mentioned above by reading and performing to drawing 10 , from the record medium is made to

realize. In addition, when realizing the function of the toy mentioned above using the personal computer Since an image processing to which a beak, feather, etc. of a parakeet on a screen are moved is performed when displaying the image of a parakeet on a monitor 51 and generating voice from a loudspeaker 20 as a substitute sewing [ which is shown in drawing 1 ]-based [ 1 ], as already stated The image processing mentioned above instead of the processing which drives the driving gear shown in step Sa24 of drawing 3 will be performed.

[0118] Operation gestalt] besides [ Below, other operation gestalten of the setup 1 (setup for making the memorized specific voice utter to the specific language or the specific sound which the user emitted) in learning mode and setting mode are explained among each mode mentioned above.

[0119] With reference to [operation gestalt of everything but learning mode] drawing 11 , and drawing 12 , other gestalten of the learning mode shown in drawing 3 are explained. When a toy learns one language, by the time it ends from initiation of study, two or more phases (study phase) will be established, and a user enables it to distinguish in which phase the language under present study is in the learning mode in this operation gestalt. In addition, in this operation gestalt, the description pattern of the language corresponding to the identifier the toy concerned was named at least shall be beforehand memorized by the description pattern storage section 25.

[0120] First, the electric power switch of the toy of this operation gestalt is turned on, and if a learning mode carbon button is pushed when a control section 29 changes into the depression standby condition of the mode selection carbon button of a control unit 3, a control section 29 will start processing of learning mode.

[0121] First, it progresses to step Sf1 of drawing 11 , voice registration processing is started, and, subsequently it judges whether voice input occurs from a microphone 10 at step Sf2. Here, when voice is not inputted, a decision result serves as No, step Sf2 is processed again, and it will be in a voice input standby condition hereafter until language (or sound) is inputted. And when voice is inputted into a microphone 10 and a detecting signal is detected from the voice input detecting element 22, a decision result serves as Yes, it progresses to step Sf3, and the voice data outputted from A/D converter 13 is stored temporarily in the voice data modification section 14.

[0122] Next, the value of the random number which progressed to step Sf4, generated the random number, and was generated at step Sf5 is compared with the predetermined value alpha. When the value of a random number is beyond the predetermined value alpha, a decision result serves as No and it returns to step Sf1. Moreover, when the value of a random number is smaller than the predetermined value alpha, a decision result serves as Yes, it progresses to step



Sf6, and voice inputted after this is learned.

[0123] Here, processing of steps Sf1-Sf5 mentioned above will not surely start study of language (or sound), if a learning mode carbon button is pushed in order to make the toy in this operation gestalt simulate the parakeet more like a genuine article, but based on the random number, since the case where study is not started is produced, it is prepared.

[0124] If it shifts to step Sf6, a control section 29 will register the audio description pattern inputted at step Sf2 to the description pattern storage section 25. Moreover, the voice data stored temporarily in the voice data modification section 14 with this is memorized in the voice data storage section 15 as "language under study." Next, it progresses to step Sf7, the phase parameter STEP corresponding to the voice data memorized as "language of study", the count register IN of an input, and the storage maintenance parameter MH are generated, respectively, and the value of each parameter and a register is set to 0, respectively.

[0125] Here, the phase parameter STEP is a parameter which shows the study phase of "the language under study", and the value of the integer of 0-3 is taken, and it means that even 0-2 are the language under study, and means that 3 is language [ finishing / study ]. Moreover, actuation sewing-based [ according to the value of this parameter / 1 ] is set up beforehand. That is, for example, in

addition to the actuation at the time of 2, actuation of generating the cry of a parakeet is beforehand set up at the time of 3 which the value of the phase parameter STEP makes flap an aerofoil while the time of 2 which the time of 1 which the time of 0 makes open and close a beak once makes flap an aerofoil makes a beak repeat, open and close.

[0126] Moreover, the count register IN of an input is a register with which the count of an input of the language inputted after step Sf6 is memorized. Furthermore, the storage maintenance parameter MH is a parameter for determining whether eliminate the voice data memorized by the voice data storage section. In addition, about the detail of the storage maintenance parameter MH, it mentions later.

[0127] And after generating the parameter and register which were mentioned above, sewing-basis [ 1 ] is made to perform actuation according to the value of the phase parameter STEP. And it progresses to step Sf8 and speech recognition is started. That is, if it progresses to step Sf9, it will be in a voice input standby condition like step Sf2 and voice is inputted into a microphone 10, it progresses to step Sf10, and the sequential comparison of the description pattern of the voice outputted from the description pattern extract section 21 and each description pattern registered into the description pattern storage section 25 will be carried out, and inputted voice will be recognized.

[0128] Next, it is judged whether the value of the random number which generated the random number at step Sf11, and was generated at step Sf12 is smaller than the predetermined value alpha mentioned above. If the value of a random number RND is beyond the predetermined value alpha, a decision result will serve as No and it will return to step Sf8. Moreover, if the value of a random number is under the predetermined value alpha, a decision result will serve as Yes and will progress to step Sf13 of drawing 12 .

[0129] And in step Sf13, the voice recognized at step Sf10 judges whether it is the identifier the toy of this operation gestalt was named. Here, when it is the identifier the toy was named, a decision result serves as Yes and this learning mode is completed. Moreover, when it is not the identifier the toy was named A decision result serves as No, progress to step Sf14, and it judges whether it is "the language under study" (the value of the phase parameter STEP is two or less language). When it is judged that it was not "the language under study", a decision result serves as No, makes sewing-basis [ 1 ] perform actuation according to the recognized language (step Sf15), and returns to step Sf8 after \*\*.

[0130] When it is judged at step Sf14 here that it is not "the language under study" [ whether the recognized language (or sound) is language / finishing / study / (the value of the phase parameter STEP is 3 language), and ] It is one

case of whether they are language other than "the language under study" (language by means of which the phase parameter STEP is not given), and a control section 29 makes sewing-basis [ 1 ] perform actuation beforehand defined about each case in step Sf15.

[0131] On the other hand, at step Sf14, when it is judged that it is "the language under study", a decision result serves as Yes, it progresses to step Sf16, and a random number is generated. And it progresses to step Sf17 and the value of a random number judges whether it is under the predetermined value beta. Here, the predetermined value beta is a numeric value established in order to determine whether advance one step of study phases.

[0132] And when the value of a random number is under the predetermined value beta, while a decision result serves as Yes, progresses to step Sf18 and increments one value of the phase parameter STEP, a predetermined fixed value (for example, 10) is added to the value of the storage maintenance parameter MH. Moreover, the value of the count register IN of an input is reset to 0. It means that one step of study phases of the voice data memorized as "language under study" was advanced to the voice data storage section 15 by this processing.

[0133] Moreover, in step Sf17 mentioned above, when the value of a random number is beyond the predetermined value beta, a decision result serves as No

and shifts to step Sf20. And it judges whether one value of the count register IN of an input was incremented, and the value of the count register IN of an input became larger than the predetermined value gamma at step Sf21. If the value of the count register IN of an input is larger than the predetermined value gamma, a decision result serves as Yes and it progresses to step Sf18, and the value of the phase parameter STEP will be incremented one and one step of study phases of the recognized "the language under study" will be advanced. Thus, when the same "language under study" has been recognized more than a certain count by processing of step Sf21, one step of study phases is advanced compulsorily.

[0134] Moreover, after one step of study phases is advanced at step Sf18, it judges whether it progressed to step Sf19 and the value of the phase parameter STEP was set to 3. With [ the value of the phase parameter STEP ] two [ or less ], it is regarded as what a decision result serves as No and study of the recognized "the language under study" has not ended, and returns to step Sf8.

[0135] On the other hand, if the value of the count register IN of an input is below the predetermined value gamma in step Sf21, a seal result will serve as No and sewing-basis [ 1 ] will be made to perform actuation according to the value of the phase parameter STEP at step Sf22. Thereby, it can recognize in which phase a user has the voice recognized at step Sf10 now. And after processing of step

Sf22 is completed, it returns to step Sf8 of drawing 11 .

[0136] Henceforth, when it carries out by repeating processing of steps Sf1-Sf22 and the value of the phase parameter STEP is set to 3 at step Sf18, the decision result of step Sf19 serves as Yes, it progresses to step Sf23, and sewing-basis [ 1 ] is made to perform actuation corresponding to the value (= 3) of the phase parameter STEP. Thereby, a user can recognize what one "the language under study" was learned for. And it progresses to step Sf24, the value of the storage maintenance parameter MH is set to 40, and learning mode is ended.

[0137] in addition, when one decrement of the value of the storage maintenance parameter MH mentioned above is carried out whenever it carried out one increment only once however, one day -- and passed on the 1st, when the voice corresponding to the storage maintenance parameter MH had been recognized with other modes, for example, execute mode etc., and a value is set to 0, the voice data \*\*\*\* description pattern is eliminated from each storage section. Thereby, the toy in this operation gestalt can be made to act about the low language or the low sound of input frequency, so that he may forget automatically.

[0138] Thus, in the learning mode of this operation gestalt, before shifting to processing (namely, processing after step Sf6) of substantial learning mode, as for \*\*\*\*\*, study of the language or the sound (what was memorized in the past is

included) storage (correctly storage of voice data and the description pattern) of \*\*\*\* was remembered to be by the language which is performed only once and inputted after it is performed. Moreover, since study of language is continuously performed until the identifier the toy was named is inputted or one study is completed, and actuation according to it is further performed whenever one step of study phases progresses, a user can know the progress condition of study.

[0139] [Other operation gestalten of setting mode [setting 1]], next other operation gestalten in setting mode [setup 1] are explained with reference to drawing 13 and drawing 14 . Although audio matching was performed by processing ( drawing 4 , steps Sb1, Sb3, Sb5-Sb11) in the setting mode [setup 1] explained previously by a user inputting voice, and the user itself choosing voice matching with this inputted voice first, and carrying out the depression of the decision key, audio matching is enabled with this operation gestalt, without needing such actuation.

[0140] First, if processing in setting mode [setup 1] is started, a random number will be generated in step Sg1 of drawing 13 , and then it will progress to step Sg2, and will judge whether voice data [ finishing / study ] (the value of the phase parameter STEP is the voice data of 3) is memorized to the random-number RND position [ of the voice data storage section 15 ] address. And when not memorizing, processing of steps Sg1 and Sg2 is repeated until the address with

which a decision result serves as No and voice data [ finishing / the return and study to step Sg1 ] is memorized is found.

[0141] On the other hand, at step Sg2, if the address which memorized voice data [ finishing / study ] is found, a decision result will serve as Yes and will progress to step Sg3. The value of the storage maintenance parameter MH corresponding to the voice data is read here, and a judgment whether the value of the storage maintenance parameter MH is larger than 40 is made at step Sg4. Here, when judged or less as 40, a decision result serves as No and returns to step Sg1. Moreover, when it is judged that it is larger than 40, it progresses to step Sg5, the voice data memorized to the address concerned is outputted to D/A converter 16, it is uttered from a loudspeaker 20, and a control section 29 starts the time check of predetermined time (here, it considers as for 7 seconds) at this time.

[0142] Next, it progresses to step Sg6 and a control section 29 recognizes voice inputted after this. namely, the case where it judges whether voice was inputted into the microphone 10 at step Sg7, and there is no audio input -- step Sg8 -- setting -- a time check -- it judges after initiation whether for 7 seconds has passed. Moreover, when it is judged that for 7 seconds has not passed, a decision result serves as Yes and returns to step Sg7. And if for 7 seconds passes while it carries out by repeating processing of steps Sg7 and Sg8



hereafter and voice has not been inputted at all by it, the decision result of step Sg8 will serve as No, it will progress to step Sg9, speech recognition processing will be ended, and it will return to step Sg1.

[0143] On the other hand, after starting speech recognition, if voice is inputted within 7 seconds, the decision result in step Sg7 will serve as Yes, and voice progressed and inputted into step Sg10 will be recognized. Subsequently, it progresses to step Sg11, a random number is generated, and the comparison with the predetermined value alpha is performed at step Sg12. Here, when the value of a random number is beyond the predetermined value alpha, after a decision result serves as No, progressing to step Sg13 of drawing 14 and starting a time check at step Sg6, it is judged whether for 7 seconds passed. And if for 7 seconds has not passed, a decision result serves as Yes and starts return re-degree speech recognition to step Sg6. Moreover, when having passed, a decision result serves as No and returns to step Sg1.

[0144] On the other hand, in step Sg12 of drawing 13 , when the value of a random number is under the predetermined value alpha, a decision result serves as Yes, and in step Sg14, when it is the identifier by which the recognized voice was attached to the toy of this operation gestalt, processing in setting mode [setup 1] is ended. Moreover, when it is not the identifier the toy of this operation gestalt was named, it progresses to step Sg15 and the data corresponding to the

voice recognized at step Sg10, i.e., the value of the storage parameter MH, are read.

[0145] and the time check of predetermined time [ in / when the value of the storage parameter MH progressed and read to step Sg16 is 40 or less, a decision result serves as No and / at step Sg17 / step Sg6 ] -- it is judged after initiation whether for 7 seconds passed, when for 7 seconds has not passed, it returns to step Sg6, and return and when having passed, it returns to step Sg1.

[0146] On the other hand, at step Sg16, when the value of the read storage parameter MH is larger than 40, a decision result serves as Yes, and it progresses to step Sg18, and judges whether there is associated data RL to the voice data (henceforth matched voice data) uttered at step Sg5.

[0147] Associated data RL is set in setting mode [setup 1] here. To matched voice data The information on the address that are data generated to this recognized voice, and matched voice data is memorized in case the voice recognized at step Sg10 is matched for the first time, It consists of a master flag (initial value 0) which shows whether they are the numeric value which shows degree of association with matched voice data, and the no by which matching with matched voice data is made. Moreover, when the value of this data is over 120, 1 is set to the master flag mentioned above as that by which matching with matched voice data is made (it mentions later).

[0148] And in step Sg18, since the associated data RL to matched voice data does not exist when matching is performed for the first time, a decision result serves as Yes, it progresses to SUTEBBU Sg19, and the associated data RL to matched voice data is generated. Here, the value of associated data RL generates a random number, and let it be the value which applied the value which subtracted 10 from the predetermined value beta ( drawing 12 , step Sf17 reference) mentioned above by the random numbers.

[0149] In addition, when the associated data RL generated in SUTEBBU Sg19 is memorized with the storage parameter MH after that and the with beam voice corresponding to the above has been recognized at step Sg10 next time, in step Sg15, the value shall be read with the value of the storage parameter MH. At this time, the value of the associated data RL which has the address information of the voice data uttered at step Sg5 shall be read with reference to the address of the matched voice data contained in each among each associated data RL memorized to the voice recognized at step Sg10.

[0150] Moreover, in step Sg18, when associated data RL has already existed, a decision result serves as No, it progresses to step Sg20, and 1 is added to the value of associated data RL. And after processing of SUTEBBU Sg19 or Sg20 is completed, it progresses to step Sg21 and judges whether the value of associated data RL is over 120. If the value of associated data RL is 120 or less,

a decision result serves as No, and it will return to step Sg1 noting that the voice recognized at step Sg10 is not matched to the voice data pronounced at step Sg5.

[0151] On the other hand, a decision result serves as Yes, progresses to step Sg22, and sets the master flag of associated data RL to 1 noting that both are matched, when the value of associated data RL is larger than 120. And the language (language pronounced at step Sg5) which progressed to step Sg23 and was matched with it, and the matched language (language recognized at step Sg10) are made to pronounce by a unit of 5 times by turns, and it notifies that both were matched to the user.

[0152] Thus, if processing of setup 1 is performed among the setting modes mentioned above If sequential utterance of the voice data [ finishing / study ] is carried out at intervals of 7 seconds in the voice data storage section 15 in the toy of this application operation gestalt and a user inputs voice in the meantime If the associated data RL in which degree of association with the voice data uttered just before it memorized corresponding to this inputted voice data and the above-mentioned user inputted voice is shown is read and the value of that parameter exceeds 120 As that with which both were matched, both language (or sound) is uttered by a unit of 5 times by turns.

[0153] moreover, the case where the contents set up by [setup 1] are performed

in execute mode -- for example, -- "-- now -- " -- \*\*, if the language to say is recognized "-- now -- " -- \*\* -- among each associated data RL memorized to the language to say The address information contained in the associated data RL with which the value of a master flag is 1 is referred to, and the voice data memorized to the address is read from the voice data storage section 15 one by one, and is generated from a loudspeaker 20.

[0154] In addition, when one decrement of the value of the associated data RL mentioned above is carried out for example, whenever it passed on the 1st, and the value turns into a predetermined value (for example, 0), the value of a master flag is reset to 0 and you may make it cancel matching. Thereby, it can be made to act, if correspondence of language always is not trained to the toy in this operation gestalt as if he forgot automatically correspondence of the language made to memorize.

[0155] Furthermore, the function in other operation gestalten mentioned above may be realized like the operation gestalt explained previously on a personal computer as shown in drawing 10 , and the so-called TV-game special-purpose machine.

[0156]

[Effect of the Invention] As explained above, while memorizing the sound signal corresponding to the uttered voice according to invention given in claims 1, 15,

and 17 Recognize the voice, and since the sound signal corresponding to the voice which could set up correspondence with the sound signal remembered to be the recognized voice, and has been recognized based on the setup is changed and outputted to voice The voice for which it not only carries out [ voice / which was only inputted ] parrotry generating as it is, but it asks to the uttered voice can be made to be able to utter, and the voice which therefore answers to the uttered voice can be diversified. Moreover, the toy which cannot get bored easily can be offered by preparing in the interior or the exterior of sewing-basis which has especially a doll or the configuration of an animal, and indoor decoration.

[0157] Moreover, since according to invention according to claim 2 utterance from a toy is continued until it repeats and utters the voice beforehand specified that the time of day which the user set up comes and a user utters similarly the voice specified beforehand, this toy can be used as an alarm clock.

[0158] Moreover, if the voice which the user set up is uttered according to invention according to claim 3, since current time will be notified by voice, it is not necessary to turn a look to the dial face of a clock in detail to know time of day. Moreover, since voice does not notify current time when the sound or voice except the user having set up is uttered, a possibility that the sound which users, such as a noise, do not mean, for example may notify current time accidentally

disappears.

[0159] Moreover, since voice notifies current time at the time of day which the user set up according to invention according to claim 4, it not only can use it as a usual timer, but it can know the present time of day.

[0160] Moreover, if the time of day which the user set up comes according to invention according to claim 5, since the voice similarly specified by the user will be uttered, the contents of the action which a user should take at predetermined time of day etc. can be notified with voice, and a user can reconfirm his schedule.

[0161] Moreover, it is lost that a uniform reaction is shown, and it can be made to act, since the voice which memorized the inputted voice to arbitration and was memorized to this arbitration is uttered to arbitration according to invention according to claim 6 as if it had reacted with autonomous volition. For this reason, there is unexpected nature and the toy which cannot get bored easily can be realized.

[0162] Moreover, rather than that it remembers simply the voice which faced memorizing exact voice and was inputted to be as it is since it is influenced with a random number whether the voice which the user uttered is memorized correctly according to invention according to claim 7, simulation of the state which masters \*\*\*\* of the animal which masters \*\*\*\* can be carried out, and unexpected nature can offer increase and the toy which cannot get bored easily.

[0163] Moreover, since the voice which the user uttered is uttered by different tone quality according to invention according to claim 8, simulation of the animal which masters \*\*\*\* can be carried out, and the insipidness that oneself hears the voice which the user uttered can be canceled.

[0164] Moreover, when the voice uttered from the same toy as the toy by this invention is inputted according to invention according to claim 9 If it sets up so that the numerical range of the generated random number may be narrowed in such a case, since the numerical range of the generated random number changes When the rate of memorizing voice correctly when making voice memorizing with toys memorizes each-other voice and suits with increase and the same toys, distinction-ization which can memorize exact voice easily is made and the way different from the case where a user makes his voice memorize again of enjoying oneself becomes possible.

[0165] Moreover, since according to invention according to claim 10 the numerical range of the generated random number changes when the same voice is inputted more than the count of predetermined, if it sets up, for example so that the numerical range of the generated random number may be narrowed when such, the phenomenon in which it does not utter as it inputted very much, even if it inputted the voice same how many times is avoidable.

[0166] Moreover, since the numerical range of the generated random number



changes when the value by which counting was carried out with the free run counter which always performs counting is in agreement with a predetermined value according to invention given in claims 11 and 12, in making exact voice memorize, gamble nature can be considered, and a user's interest can be lengthened. Moreover, since predetermined voice informs a user of that when the numerical range of the generated random number changes, a user's interest can be fueled.

[0167] Moreover, since sequential generating of the memorized voice can be carried out altogether according to invention according to claim 13, while the check of the memorized voice is attained, when the message addressed to the partner is memorized, the message can be certainly told to a partner.

[0168] Moreover, since according to invention according to claim 14 some bodies of a sewing basis toy which have a doll or the configuration of an animal move in case a sound or voice occurs, if it is made for a beak etc. to move for example, when sewing-basis of a parakeet etc. is adopted as the toy of this invention, effectiveness as if the real parakeet had talked can be given.

[0169] Moreover, since the image of the displayed man or animal changes in case according to invention given in claims 16 and 18 the image of a man or an animal is displayed on an image display device and a sound or voice is generated, also visually, it can be enjoyed.

[0170] Moreover, according to invention according to claim 19, the voice data outputted from the voice data output means is stored temporarily. A random number is generated when the same voice data as the this voice data stored temporarily is outputted from a voice data output means. When the size relation between this random number and a predetermined value fulfills predetermined conditions the number of predetermined times Since change which is different, respectively in the image of those who make it display on an image display device, or an animal arises whenever it memorizes as voice data matched with the voice one voice data has been recognized to be and fulfills the above-mentioned predetermined conditions The language or the sound which the user is going to make it memorize can distinguish in which phase it is in the process memorized as voice data matched with the recognized voice.

[0171] Moreover, in case voice data is memorized as voice data matched with the recognized voice according to invention according to claim 20 Since the voice data concerned is eliminated when the storage retention data for determining whether eliminate the voice data concerned to the voice data concerned occurred, these storage retention data decrease whenever predetermined time passed, and it becomes below constant value Those who make it display on an image display device, or an animal can make it act about the low language or the low sound of input frequency as if it forgot automatically.

[0172] Moreover, the voice data which was memorized as voice data matched according to invention according to claim 21, In case the recognized voice is made to be matched, while the voice data memorized as voice data matched is outputted to a voice output means one by one at intervals of predetermined time Recognize the voice inputted in the meantime and the information which shows degree of association with the voice data outputted to the voice output means just before that occurs to this recognized voice. The voice which carried out [ above-mentioned ] recognition next time immediately after outputting the voice data in which relation is shown using the information generated to the recognized voice to a voice output means Since the voice data made to memorize as voice data matched with the voice which carried out [ above-mentioned ] recognition the account of a top is matched when have been recognized again, and the degree of association which the information shows is made to increase and this degree of association fulfills predetermined conditions Voice can be matched without needing actuation of the user itself choosing voice matching with a certain voice, and carrying out the depression of the decision key.

[0173] Furthermore, since correspondence with the matched voice and voice data is canceled when according to invention according to claim 22 it decreases whenever the degree of association which the above-mentioned information

shows carried out predetermined time progress, and this degree of association becomes below constant value. It can be made to act, if correspondence of language always is not trained to those who make it display on an image display device, or an animal as if he forgot automatically correspondence of the language made to memorize.

---

## DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1] It is the front view showing the appearance of the toy equipped with the speech recognition and the speech generation device by 1 operation gestalt of this invention.

[Drawing 2] It is the block diagram showing same sound voice recognition and the configuration of a speech generation device.

[Drawing 3] It is the flow chart which shows the procedure of the learning mode in same sound voice recognition and a speech generation device.

[Drawing 4] It is the flow chart which shows the procedure in the setting mode in same sound voice recognition and a speech generation device.

[Drawing 5] It is the flow chart which shows the procedure in the setting mode in

same sound voice recognition and a speech generation device.

[Drawing 6] It is the flow chart which shows the procedure of the execute mode in same sound voice recognition and a speech generation device.

[Drawing 7] It is the flow chart which shows the procedure of the execute mode in same sound voice recognition and a speech generation device.

[Drawing 8] It is the flow chart which shows the procedure of the random mode in same sound voice recognition and a speech generation device.

[Drawing 9] The procedure in the message mode in same sound voice recognition and a speech generation device is shown.

[Drawing 10] It is the block diagram showing the configuration at the time of using a personal computer in same sound voice recognition and a speech generation device.

[Drawing 11] It is the flow chart which shows the procedure of other operation gestalten of the learning mode in same sound voice recognition and a speech generation device.

[Drawing 12] It is the flow chart which shows the procedure of other operation gestalten of the learning mode in same sound voice recognition and a speech generation device.

[Drawing 13] It is the flow chart which shows the procedure of other operation gestalten of same sound voice recognition, and the setting mode and a setup 1

in a speech generation device.

[Drawing 14] It is the flow chart which shows the procedure of other operation gestalten of same sound voice recognition, and the setting mode and a setup 1 in a speech generation device.

[Description of Notations]

1 Sewing-basis

1a Beak

2 Body

3 Control Unit

4 Display

10 Microphone

11 18 Amplifying circuit

12 Filter

13 A/D Converter

14 Voice Data Modification Section

15 Voice Data Storage Section

16 D/A Converter

17 Voice Conversion Circuit

19 Sound-Volume Equalization Circuit

20 Loudspeaker

21 The Description Pattern Extract Section

22 Voice Input Detecting Element

23 Timer

24 The Toy Description Pattern Storage Section

25 The Description Pattern Storage Section

26 Counter

27 Random-Number-Generation Section

28 Driving Gear

29 Control Section

50 Personal Computer

51 Monitor

52 Keyboard

53 Sound-Source Board